



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

Réseaux, Télécommunications, Systèmes et Architecture

Présentée et soutenue par :

M. CHAKADKIT THAENCHAIKUN

le mardi 22 novembre 2016

Titre :

ECONOMIE D'ENERGIE EN RESEAU FILAIRE: INGENIERIE DE TRAFIC
ET MISE EN VEILLE

Ecole doctorale :

Mathématiques, Informatique, Télécommunications de Toulouse (MITT)

Unité de recherche :

Institut de Recherche en Informatique de Toulouse (I.R.I.T.)

Directeur(s) de Thèse :

MME BEATRICE PAILLASSA

Rapporteurs :

M. LAURENT LEFEVRE, ECOLE NORMALE SUPERIEURE DE LYON

Mme ISABELLE CHRISMENT, UNIVERSITE DE LORRAINE

Membre(s) du jury :

M. CHRISTOPHE CHASSOT, INSA TOULOUSE, Président

M. EMMANUEL DUBOIS, CENTRE NATIONAL D'ETUDES SPATIALES CNES, Membre

M. GENTIAN JAKLLARI, INP TOULOUSE, Membre

Mme BEATRICE PAILLASSA, INP TOULOUSE, Membre

Résumé

Les travaux portent sur l'économie d'énergie dans le secteur des technologies de la communication et plus particulièrement dans les réseaux filaires. La technologie support de nos travaux est Ethernet qui historiquement utilisée dans les entreprises est actuellement déployée dans les réseaux d'accès et de coeur des opérateurs. Notre objectif est d'économiser de l'énergie par une mise en sommeil des liens Ethernet en s'appuyant sur des mécanismes standards aisément déployables. Pour ce faire nous modélisons et évaluons le mécanisme de mise en sommeil IEEE802.3az et confrontons notre modèle à l'expérimentation sur équipements. A partir du modèle de coût obtenu, nous proposons de mettre en place une ingénierie de trafic verte fonction de la charge qui dirige le trafic dans le réseau de façon à permettre aux liens de se mettre en phase de sommeil tout en préservant une qualité de service, en évitant de créer des pertes de données dans le réseau par une concentration trop importante de trafic sur les liens. Nous distinguons plusieurs politiques d'ingénierie de complexité de mise en oeuvre différentes, que nous évaluons dans plusieurs contextes. Les résultats obtenus permettent quasiment de doubler les gains obtenus par le standard IEEE802.3az. Nous étudions l'ingénierie dans un mode distribué à partir d'un protocole de routage, OSPF, et présentons une preuve de concept dans un mode centralisé avec une architecture SDN pour laquelle nous proposons l'utilisation du routage source par segment afin de réduire le trafic de contrôle.

Abstract

This research has proposed a novel Ethernet-based global energy efficiency technology for the wired communications networks. In the development of the *green* technology, the IEEE 802.3az standard (i.e. referred to as LEGACY) was first applied to render dormant the non-utilized links to realize the local energy efficiency. To further lower the energy demand and thus achieve the global energy efficiency, this research put forward two additional algorithmic schemes (i.e. the EAGER and CARE green metrics) whereby the data are aggregated such that as many links as possible remain unutilized. The three metrics (LEGACY, EAGER and CARE) were individually applied to the distributed and centralized networks and subsequently evaluated. The experimental results showed that, by comparison, the CARE green metric outperforms the other two metrics (LEGACY and EAGER) with regard to the global energy saving in both the distributed and centralized wired networks.

Remerciements

Mes premiers remerciements iront à Béatrice PAILLASSA et Gentian JAKLLARI avec qui j'ai travaillé des mon mastère pour leur soutien permanent pendant ma thèse. J'aimerais leur adresser mes plus vifs remerciements pour tout leur dynamisme, leur tolérance, leur disponibilité, leurs pertinentes suggestions que j'ai pu apprécier tout au long de ma thèse. Ce travail n'aurait jamais pu aboutir sans leurs conseils avisés, leur soutien et la confiance qu'ils m'ont témoignés. Je souhaite leur transmettre l'expression de ma plus profonde et sincère gratitude. J'ai beaucoup appris en travaillant avec eux et je les remercie de tout mon coeur. Je voudrais leur dire que tout deux sont parmi les meilleurs professeurs que j'ai rencontrés et je pense que c'était une occasion unique pour moi de travailler avec eux.

Mes très vifs remerciements s'adressent aux membres du jury de thèse qui ont accepté de juger ce travail. Je suis profondément reconnaissant à mes deux rapporteurs, ISABELLE CHRISMENT et LAURENT LEFEVRE, d'avoir accepté de lire et évaluer ma thèse. Je remercie également CHRISTOPHE CHASSOT et EMMANUEL DUBOIS pour m'avoir fait l'honneur de participer a jury.

Merci au personnel du laboratoire, de l'INPT et l'IRIT pour les aides permanentes reçues. Je souhaite aussi remercier les secrétaires de l'IRIT, SAM, Sylvie et Audrey pour leur accueil, leur discussion et de nous avoir rendu la vie bien plus facile. J'adresse ma gratitude à mes amis, aux autres doctorants et tous ceux que j'ai côtoyé : Farouk, les Mohamed (Hussein et Bouazzouni), les Ahmed Khalid, Jean-Gabriel, Aziz, Abderrahim, Emilie, Patrice, Renaud, Fabien, Warodom, Kuljaree, Qiankun, Xiaoyan, Mouna, Elie, Yoann, Dorin, Sa famille de Chidkae et Xavier SERVILE,

KEEGAN, CORMON, CONQUY, KAMNERDSUPAPHON, Teeka TANTAVICHIT, Charlie KREY et bien d'autres. Je garderai en mémoire les moments magnifiques que nous avons partagés ensemble.

J'aimerais remercier également l'ensemble du personnel de l'Université de la Thaïlande, Prince of Songkla University pour pour leur aide précieuse.

Enfin et surtout, je remercie de tout coeur ma famille, ma femme, la famille de ma femme pour leur soutien sans faille. Ce grand succès n'aurait pu avoir lieu sans eux. Leurs encouragements sont les facteurs les plus importants de cet accomplissement.

Chakadkit

Table des matières

Résumé	iii
Abstract	v
Table des figures	xv
Liste des tableaux	xvii
Liste des acronymes	xix
Introduction	1
A A propos de la consommation d'énergie	1
A.1 Economie d'énergie dans les TIC	1
A.2 Le réseau filaire	3
B Présentation des travaux	5
B.1 Domaine	5
B.2 Méthodologie	5
B.3 Contributions	5
B.4 Organisation de la thèse	6
Chapitre 1 Economie d energie en réseau filaire	9
1 Panorama des technologies	10

2	Stratégies locales d'économie d'énergie	12
2.1	Adaptation de débit (Adaptation Link Rate)	12
2.1.1	Débits et économie d'énergie	13
2.1.2	Synchronisation émetteur récepteur	14
2.1.3	Politiques d'adaptation	15
2.2	Endormissement de lien	17
2.2.1	Politique de réveil	17
2.2.2	Cas du réseau optique	19
2.3	Relais de proximité	23
3	Stratégie Globale d'économie d'énergie	24
3.1	Approche trafic	25
3.1.1	Optimisation et contraintes	25
3.1.2	Ingénierie de trafic verte dans IP	27
3.1.3	Ingénierie de trafic verte dans MPLS	30
3.1.4	Ingénierie de trafic verte dans Software Defined Network	31
3.1.5	Cas des data centers	33
3.2	Approche par planification de topologie	33
3.2.1	Approche distribuée	34
3.2.2	Approche centralisée	35
4	Conclusion	36
Chapitre 2 Ethernet E cace en Énergie		39
2.1	Le standard Energy Efficient Ethernet	40
2.1.1	Intégration du mode de mise en sommeil dans l'architecture en couche	41
2.1.2	Fonctionnement	45
2.2	Implantation EEE et évaluation de consommation	47
2.2.1	Banc de test et caractéristiques de simulation	47
2.2.2	Modèle de puissance	50

2.2.3	Résultats	51
2.3	Amélioration du standard	52
2.3.1	Algorithmes de coalescence	53
2.3.2	Intérêt et problème de la coalescence de trames	56
2.4	Conclusion	58
Chapitre 3	Ingénierie verte distribuée	59
3.1	Ingénierie de trafic verte adaptée à la technologie EEE	60
3.1.1	Principaux éléments OSPF	61
3.1.2	Modèle de coût OSPF pour réseau Ethernet Efficace en Energie	62
3.1.3	Illustration des politiques d'ingénierie verte proposées	65
3.2	Evaluation du fonctionnement de l'ingénierie par simulation et expérimentation	68
3.2.1	Evaluation par simulation	69
3.2.2	Evaluation expérimentale	77
3.3	Comparaison de la proposition sur un scénario réaliste	81
3.3.1	Comparaison à une solution de mesure idéale	81
3.3.2	Comparaison à la solution optimale Green-TE	86
3.4	Architecture multicouche - Prise en compte de l'énergie lien	90
3.4.1	Proposition de métrique verte en fibre optique	90
3.4.2	Evaluation	92
3.5	Conclusion	97
Chapitre 4	Ingénierie verte centralisée par réseau programmable SDN	99
4.1	Présentation de SDN (Software Defined Network)	100
4.1.1	L'architecture de SDN	101
4.1.2	Principaux Eléments	102
4.1.3	OpenFlow	104
4.2	SDN et ingénierie de trafic	107
4.2.1	Principe du Segment routing	108

4.2.2	Partage de charge et Ingénierie verte.	110
4.3	Evaluation	112
4.4	CONCLUSION	117
Conclusion et Perspectives		119
Liste des publications		123
Bibliographie		137

Table des figures

1	Evolution de la consommation électrique totale et par catégorie d'éléments	2
2	Structure de réseau de télécommunications	3
3	Projection de trafic Internet en filaire et sans fil	4
1.1	Approche par stratégie globale versus stratégie locale des techniques d'économie d'énergie dans les réseaux filaires	12
1.2	Principe de l'adaptation de débit	13
1.3	Evaluation de la puissance consommée selon le débit en Ethernet	14
1.4	Principe du fonctionnement de l'ADSL avec économie d'énergie	17
1.5	Approche simple à deux états de l'endormissement de lien.	18
1.6	Illustration de l'endormissement en réseau optique d'accès XG-PON.	21
1.7	Comparaison des méthodes d'endormissement optique	22
1.8	Illustration de la méthode de sommeil avec relais de proximité.	23
1.9	Historique de l'ingénierie	26
1.10	Exemple de calcul de poids pour un routage à énergie selon 3 périodes de la journée : a) matin, b) après-midi c) soir	28
1.11	Fonctionnement du routage HotPLUZ	29
1.12	Reroutage de trafic sur LSP économe en énergie (e-LSP)	30
1.13	Influence du nombre de règles SDN sur le routage à économie d'énergie	32
1.14	Exemple de topologie en réseau DCN	34
2.1	Interface entre le client LPI et la couche de réconciliation de Ethernet	42

2.2	Architecture standard des blocs fonctionnels de la couche physique Ethernet	44
2.3	Mode basse consommation Ethernet standardisé (LPI : Low Power Iddle)	45
2.4	Machine à état finis de la fonctionnalité de mise en sommeil Ethernet Standard	46
2.5	Banc d'essai IEEE802.3az	47
2.6	Architecture de mesure	49
2.7	Comparaison de la consommation d'énergie IEEE802.3az entre banc d'essai et simulation	52
2.8	Mécanisme de coalescence	55
2.9	Fonctionnement du standard EEE avec coalescence	55
2.10	Consommation d'énergie de l'IEEE 802.3az avec et sans coalescence	57
3.1	Economie d'énergie du standard IEEE 802.3az à différentes charges de trafic	63
3.2	Représentation graphique de EAGER	64
3.3	Représentation graphique de CARE.	65
3.4	Illustration de OSPF-EE avec la métrique EAGER	66
3.5	Illustration de CARE-50.	67
3.6	Topologie de 10 routeurs générée par BRITE	69
3.7	Illustration du processus de génération de trafic.	71
3.8	Comparaison sur une topologie modèle Barabasi des politiques d'ingénierie verte	72
3.9	Utilisation des liens-Topologie Modèle Barabasi- a) 100Mbps b) 500Mbps c) 700Mbps	72
3.10	Délai de bout en bout- a) 100Mbps b) 500Mbps c) 700Mbps	73
3.11	Zone d'économie d'énergie-Cas d'un trafic d'entrée à 700 Mbps	75
3.12	Eléments du banc de test	77
3.13	Configuration du banc de test	78
3.14	Configuration pour LEGACY OSPF	79
3.15	Configuration pour EAGER et CARE	80
3.16	Pourcentage d'économie d'énergie globale du réseau de test.	80
3.17	Topologie du réseau Abilene	83
3.18	Economie d'énergie globale sur le réseau Abilene.	84
3.19	Maximum link utilization du réseau Abilene mesuré toutes les 5 minutes	85

3.20	Compromis entre MLU et économie d'énergie pour la topologie du réseau Abilene . . .	85
3.21	Cumulative distribution functions (CDF) de l'utilisation maximale des liens a) en heures creuses (b) en heures pleines	86
3.22	Comparaison de performances Green TE, EAGER, CARE	89
3.23	Utilisation des liens	89
3.24	Lien en support optique	90
3.25	Illustration du facteur Amplificateur (AF)	92
3.26	Distance des liens dans Abilene	94
3.27	Résultats d'évaluation	95
4.1	Architecture de réseau SDN	101
4.2	Principaux Eléments SDN	102
4.3	Architecture OpenFlow	104
4.4	Échange de messages Openflow	106
4.5	Opérations du segment routing	109
4.6	Empilement de labels en segment routing	109
4.7	Illustration de partage de charge dans ECMP	110
4.8	Architecture de l'ingénierie verte en SDN et segment routing	112
4.9	Topologie expérimentale SDN-Abilene	113
4.10	Economie d'énergie sur réseau SDN-Abilene en segment routing (a) tunnel, (b) partage de charge	115
4.11	CDF de l'utilisation maximale de liaison sur réseau SDN-Abilene en segment routing (a) tunnel, (b) partage de charge	116
4.12	CDF de délai de bout en bout sur réseau SDN-Abilene en segment routing (a) tunnel, (b) partage de charge	117

Liste des tableaux

1.1	Panorama des technologies en économie d'énergie réseaux	11
2.1	Paramètres du mode LPI par version de débit Ethernet	45
2.2	Paramètres de consommation en transmission	50
2.3	Paramètres de consommation en émission	50
2.4	Paramètres de simulations	57
3.1	Synthèse des performances	76
3.2	Valeurs des niveaux d'économie d'énergie globale et de délai de bout en bout	76
3.3	Consommation d'énergie pour équipements en réseau optique multicouches.	94

Liste des acronymes

ADSL Asymmetric Digital Subscriber Line.

ADSL2 Asymmetric Digital Subscriber Line 2.

AF Amplifier Factor.

API Application Programming Interface.

ARP Address Resolution Protocol.

ATM Asynchronous Transfer Mode.

BGP Border Gateway Protocol.

BRITE Boston university Representative Internet Topology gEnerator.

CARE Congestion aware Augmented eneRgy efficient Ethernet.

CISCO VNI CISCO Visual Networking Index.

CLI Command Line Interface.

CMOS Complementary metal-oxide-semiconductor.

CoS Class of Service.

CR-LDP Constraint-based Routing Label Distribution Protocol.

CRS-3 Carrier Routing System 3.

D-ITG Distributed Internet Traffic Generator.

DCN Data Center Network.

DOCSIS Data Over Cable Service Interface Specification.

DSL Digital Subscriber Line.

DSLAM Digital Subscriber Line Access Multiplexer.

DTP Dual-Threshold Policies.

EAGER Extremely AuGmented Energy efficient etheRnet.

EAR Energy Aware Routing.

ECMP Equal Cost Multi Path.

ECONET low Energy CONsumption NETworks.

EEE Energy Efficient Ethernet.

EPON Ethernet Passive Optical Networks.

GPON Gigabit-capable Passive Optical Networks.

IBGP Internal Border Gateway Protocol.

IETF Internet Engineering Task Force.

IP Internet Protocol.

ITU International Telecommunication Union.

LLDP Link Layer Discovery Protocol.

LPI Low Power Idle.

LSA Link State Advertisements.

LSDB Link State Database.

MAC Media Access Control.

MCF Multi Commodity Flow.

MILP Mixte Integer Linear Programming.

MIP Mixed Integer Programming.

MLU Maximum Link Utilization.

MPLS Multi-Protocol Label Switching.

MPLS-TP Multiprotocol Label Switching - Transport Profile.

NS3 Network Simulator 3.

OAM Operation And Maintenance.

OLT Optical Line Terminal.

ONF Open Network Foundation.

ONOS Open Network Operating System.

ONU Optical Network Unit.

OSPF Open Shortest Path First.

OTN Optical Transport Network.

PCS Physical Coding Sublayer.

PLAN Packet Language for Active Networks.

PMA Physical Medium Attachment Sublayer.

PMD Physical Medium Dependant.

R-EAR Reliable Energy-Aware-Routing.

RCP Routing Control Platform.

ROD Routing On Demand.

RSVP-TE Resource ReSerVation Protocol - Traffic Engineering.

SDN Software-Defined Networking.

SID Segment Identifiers.

SNMP Simple Network Management Protocol.

STP Spanning Tree Protocol.

TCAM Ternary Content Addressable Memory.

TCP Transmission Control Protocol.

TE Traffic Engineering.

TIC technologies de l'information et de la communication.

TM Traffic Matrix.

UDP User Datagram Protocol.

VDSL2 Very High Speed Digital Subscriber Line 2.

VLAN Virtual LAN.

WDM Wavelength Division Multiplexing.

X-GPON 10-Gigabit-capable Passive Optical Network.

Introduction

Nous explicitons le problème d'économie d'énergie traité dans la thèse en situant plus spécifiquement l'économie d'énergie en réseau filaire puis nous présentons le contexte de nos travaux

A A propos de la consommation d énergie

Nous présentons en premier lieu le problème que nous adressons dans cette thèse, l'économie d'énergie en réseau filaire, en situant la thématique dans un cadre plus global d'économie d'énergie qui est celui des [technologies de l'information et de la communication \(TIC\)](#). Nous précisons la part du composant réseau dans la consommation d'énergie des [TIC](#).

A.1 Economie d énergie dans les TIC

Le green networking ou plus précisément réseau vert en français s'inscrit dans les approches environnementales de protection et de valorisation des ressources. Le réseau filaire vert est un réseau économe en énergie qui dispose de suffisamment d'énergie pour fonctionner avec cependant le souci de diminuer sa consommation comme le préconise de nombreuses directives mondiales. Ainsi en Octobre 2014, les dirigeants européens ont exprimé leur souhait de poursuivre leurs efforts d'économie pour la période 2021 à 2030 dans le but de réduire les émissions de CO₂ en 2030 de 30% par rapport à 2005. En 2005, la consommation d'énergie dans le domaine des [TIC](#) avait été estimée à elle seule à 2% des émissions de CO₂.

Pour la période 2007-2012, l'étude effectuée par [\[VHLL⁺14\]](#) estime la croissance annuelle du domaine [TIC](#) en séparant 3 catégories : les ordinateurs personnels, les data centers et les réseaux de

communication. L'étude indique une croissance de consommation plus rapide dans chaque catégorie que celle de la totalité de la croissance de consommation électrique mondiale. Avec 3% pour l'augmentation globale et 10% pour la catégorie réseau, 5% pour les ordinateurs personnels et 4% pour les data centers. Il est à noter que l'augmentation globale était d'environ de 3,9 % en 2007 et qu'elle atteint 4,6% en 2012. Quant à la consommation de chaque catégorie d'éléments étudiés, elle est quasi équivalente comme l'indique la figure 1.

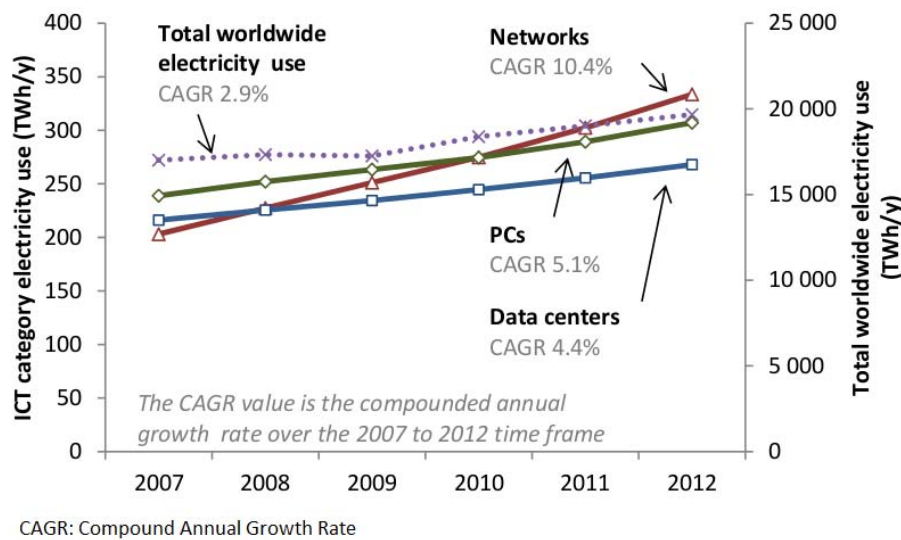


FIGURE 1 – Evolution de la consommation électrique totale et par catégorie d'éléments [VHLL⁺14].

Le défi de recherche auquel doivent faire face les communautés académique et industrielle est de diminuer la consommation d'énergie dans les différentes catégories des éléments. L'objectif ambitieux énoncé par le consortium GreenTouch en 2015 est d'arriver à 98% de réduction d'énergie en 2020 par rapport au scénario de référence élaboré antérieurement (en 2010) [Gre15]. Cet objectif sera atteint grâce à la conjonction de travaux dans les domaines des technologies des composants, architectures, algorithmes et protocoles.

A.2 Le réseau laire

Si nous détaillons plus avant la catégorie d'éléments dite "réseau", nous relevons une grande variété de moyens pour diminuer la consommation qui portent sur différents segments de la structure du réseau avec des déclinaisons technologiques variées (technologies, 2G, 3G, 4G, optique, [Asymmetric Digital Subscriber Line \(ADSL\)](#), [Internet Protocol \(IP\)](#), [Multi-Protocol Label Switching \(MPLS\)](#)...) relevant du domaine filaire et du domaine sans fil comme illustré sur la figure 2.

L'architecture présentée dans cette figure est typique des réseaux de télécommunications. C'est une architecture hiérarchique à différents niveaux.

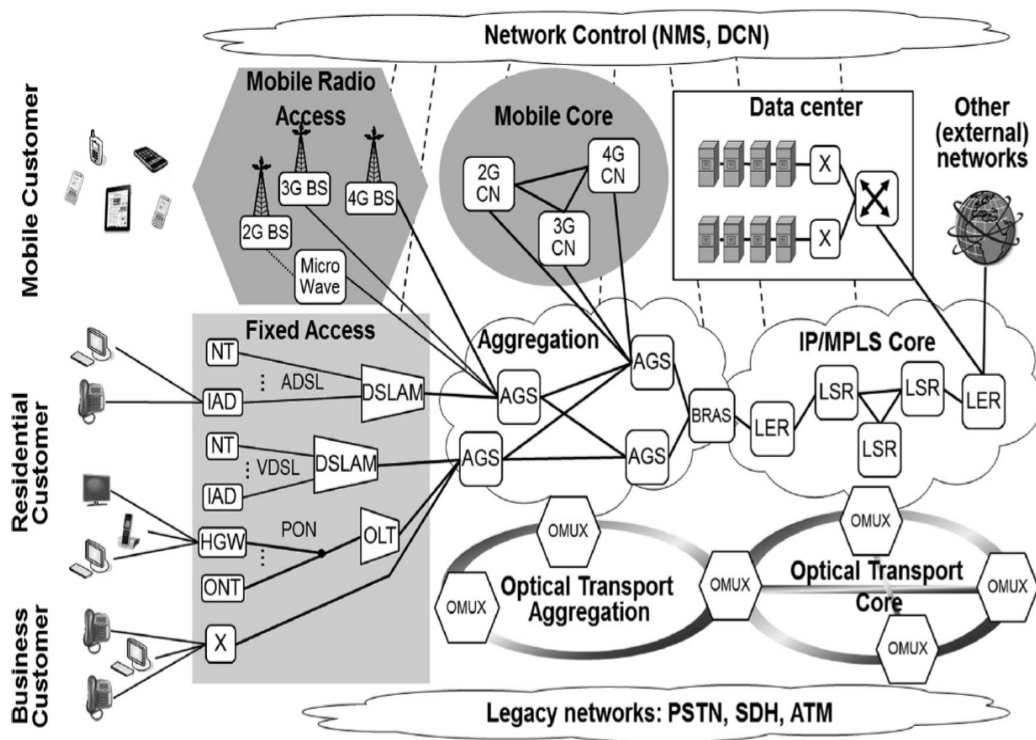


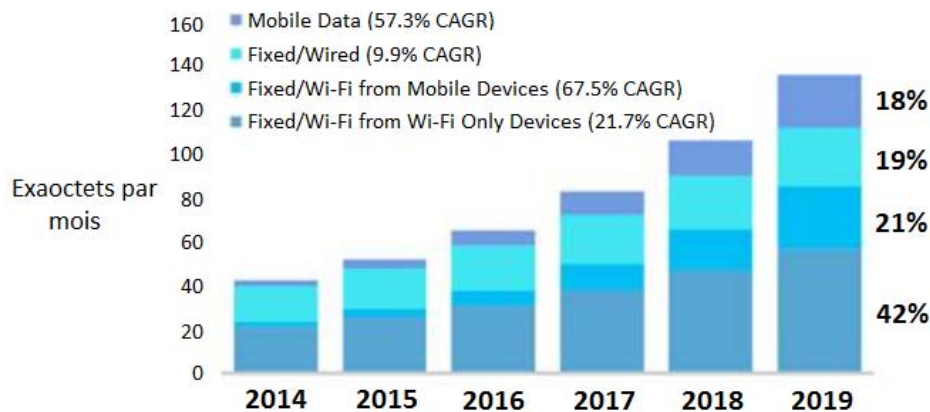
FIGURE 2 – Structure de réseau de télécommunications [LKB⁺14]

- Réseau d'accès : raccorde les utilisateurs sur les segments locaux. Le réseau d'accès est majoritairement filaire pour les utilisateurs résidentiels, les entreprises, et, il est sans fil pour les utilisateurs mobiles.

- Réseau d'agrégation : regroupe les données issues des réseaux d'accès. Sa déclinaison peut être en réseau filaire ou réseau sans fil.
- Réseau de coeur : c'est l'épine dorsale de l'architecture. Sa fonction principale est d'acheminer le plus rapidement possible les informations d'un réseau à un autre réseau. C'est majoritairement un réseau filaire hautement fiable qui s'appuie sur des technologies optiques avec des composants redondants et hautement fiable.

Le défi de recherche se complexifie avec les prévisions de croissance des besoins de communication. En effet les prédictions de trafic entre 2010 et 2020 indiquent un accroissement très important des besoins de transfert comme nous pouvons le voir avec les données de [CISCO Visual Networking Index \(CISCO VNI\)](#) en Figure 3 [cis15].

Les prévisions de croissance de trafic apparaissent dans tous les segments de la communication : le mobile mais également le filaire que ce soit seul en ou en combinaison avec le sans fil.



Source: Cisco VNI Global IP Traffic Forecast 2014-2019
CAGR: Compound Annual Growth Rate

FIGURE 3 – Projection de trafic Internet en filaire et sans fil [cis15].

B Présentation des travaux

B.1 Domaine

Les travaux présentés dans cette thèse portent sur les économies d'énergie en réseaux filaires. Secteur non négligeable d'économie d'énergie qui d'après l'étude du projet Européen ECONET ([low Energy CONsumption NETworks \(ECONET\)](#)) [BDB⁺11] permettra grâce à la combinaison de différentes technologies un gain de 58% pour le réseau de coeur et jusqu'à 70% pour le réseau d'accès. Nous nous intéressons au gain d'énergie apporté par des mécanismes protocolaires indépendants des applications.

La technologie support de nos travaux est la technologie Ethernet qui a l'origine a été définie pour les réseaux d'entreprise et qui est actuellement déployée dans les réseaux d'accès et dans les réseaux de coeur. Cette technologie a fait l'objet d'études pour permettre d'en augmenter l'efficacité en terme d'énergie. Nous nous appuierons sur ces travaux pour en améliorer l'économie d'énergie.

B.2 Méthodologie

L'objectif des travaux est d'améliorer l'efficacité énergétique des transmissions d'information en s'inscrivant dans un cadre de standardisation, de façon à pouvoir déployer sans modification majeure sur les protocoles existants, nos propositions.

L'amélioration de l'efficacité énergétique est évaluée à partir de simulations de réseaux génériques résultant de modélisation de topologie et trafic, et réels, avec une modélisation de trafic découlant de traces disponibles en accès libre. L'expérimentation sur matériel conforte ces évaluations.

Nous évaluons l'amélioration apportée en regard d'une solution optimale et la situons par rapport aux travaux de la littérature.

B.3 Contributions

Les principales propositions de cette thèse sont

- L'analyse du modèle de coût en énergie de la technologie Ethernet

- La proposition de politiques d'ingénierie verte reposant sur des métriques de routage reflétant le coût en énergie, et la qualité de service.
- L'étude de mise en oeuvre dans un environnement de réseaux programmables.

B.4 Organisation de la thèse

Après la présentation du problème effectuée dans cette introduction, le document est organisé en 4 chapitres présentant : l'état de l'art et l'approche de résolution du problème que nous avons étudiée, poursuivi par une présentation et une évaluation du contexte du travail, à savoir Ethernet 802.3 az (Ethernet Econome en Energie), puis par les propositions d'ingénierie verte que nous évaluons et expérimentons avant de conclure.

Chapitre 1 : Economie d'énergie en réseau aux filaires

Ce chapitre présente un état de l'art des travaux en réseau filaires original, en ce sens que le domaine d'étude étant récent, peu de travaux ont été publiés sur ce thème couvrant tant les mécanismes standardisés que les travaux de recherche, comme nous le proposons. Nous justifions l'approche d'économie d'énergie étudiée, à savoir l'endormissement, et présentons les méthodes d'ingénierie de trafic verte que nous employons pour diminuer la consommation d'énergie en réseaux filaires.

Chapitre 2 : Ethernet Efficace en énergie

Le contexte de réseau filaire que nous étudions, Ethernet à économie d'énergie, est détaillé et analysé, via les outils de simulation et expérimentation mettant en avant l'intérêt du contexte et les améliorations de consommation énergétique qui peuvent être réalisées par les algorithmes de coalescence.

Chapitre 3 : Ingénierie verte distribuée

Nous développons des politiques d'ingénierie verte permettant d'étendre l'économie d'énergie de Ethernet par un protocole de routage distribué [Open Shortest Path First \(OSPF\)](#). Nous étudions deux métriques de complexité différentes, [Extremely AuGmented Energy efficient etheRnet \(EAGER\)](#) et [Congestion aware Augmented eneRgy efficient Ethernet \(CARE\)](#), déclinées selon le contexte cuivre

et optique et déterminons leur domaine d'intérêt selon la charge du réseau. L'étude est effectuée par simulation et comparaison à d'autres solutions.

Chapitre 4 : Ingénierie verte centralisée

Nous développons une preuve de concept de nos propositions en émulant le fonctionnement de l'ingénierie de réseau en environnement de réseau logiciel, ([Software-Defined Networking \(SDN\)](#)), approche architecturale de contrôle de réseau novatrice qui devrait être appelée à se déployer massivement chez les opérateurs de télécommunications, et proposons d'utiliser un routage par la source et par segment pour diminuer les interactions contrôleur switch.

CHAPITRE 1

Economie d energie en réseau filaire

Sommaire

1	Panorama des technologies	10
2	Stratégies locales d économie d énergie	12
2.1	Adaptation de débit (Adaptation Link Rate)	12
2.1.1	Débits et économie d'énergie	13
2.1.2	Synchronisation émetteur récepteur	14
2.1.3	Politiques d'adaptation	15
2.2	Endormissement de lien	17
2.2.1	Politique de réveil	17
2.2.2	Cas du réseau optique	19
2.3	Relais de proximité	23
3	Stratégie Globale d économie d énergie	24
3.1	Approche trafic	25
3.1.1	Optimisation et contraintes	25
3.1.2	Ingénierie de trafic verte dans IP	27
3.1.3	Ingénierie de trafic verte dans MPLS	30
3.1.4	Ingénierie de trafic verte dans Software Defined Network	31
3.1.5	Cas des data centers	33
3.2	Approche par planification de topologie	33
3.2.1	Approche distribuée	34
3.2.2	Approche centralisée	35
4	Conclusion	36

Nous explicitons dans ce premier chapitre le problème d'économie d'énergie dans les réseaux filaires traité dans cette thèse en présentant les problèmes et solutions présents dans la littérature.

Nous examinons et analysons les travaux de recherche menés pour économiser de l'énergie dans les réseaux filaires en les différenciant selon leur stratégie de déploiement, soit une stratégie locale soit une stratégie globale. Dans ce dernier cas, contrairement au premier, une coordination globale au niveau d'un domaine de déploiement est nécessaire, elle est réalisable en mode distribué ou bien en mode centralisé, avec différents niveaux de dynamicité et de connaissance de trafic.

1 Panorama des technologies

Nous avons regroupé les travaux publiés dans la littérature en quatre grandes classes d'approches comme l'indique le tableau 1.1. (Une autre classification sera trouvée en Bianzino et al. [BCRR12]).

Ces quatre approches sont :

- Adaptation de débit
- Endormissement de lien
- Utilisation d'un relais de proximité
- Topologie active

Dans la table 1.1 nous avons caractérisé les approches par un critère de répartition : local ou global.

Le critère local est partagé par les 3 premières approches. Il est possible de déployer ces technologies graduellement sur certaines entités du réseau. Ce qui présente un avantage majeur pour déployer ces méthodes. Les stratégies consistent à surveiller l'utilisation de l'entité dynamiquement et adapter sa consommation d'énergie en fonction de la performance requise. A titre d'exemple, l'entité peut passer à un état de sommeil [GS03] ou bien diminuer son débit [GCS06].

La dernière classe d'approche est quant à elle globale, elle concerne un ensemble d'équipements qui vont mettre en oeuvre une stratégie coordonnée. Le déploiement de ces stratégies se fait au niveau de l'opérateur réseau qui doit modifier le fonctionnement du réseau sur de nombreux équipements à la fois. L'objectif est de pouvoir soit endormir soit mettre en état de veille profonde certains équipements

TABLE 1.1 – Panorama des technologies en économie d'énergie réseaux

Approches	Stratégie	Fonctions et Travaux	Reference
Adaptation de débit	Stratégie Locale	Synchronisation émetteur récepteur	[GCS06][ARSG06]
		Politique d'occupation de tampons	[GCN05][NPI ⁺ 08] [GCNS08]
		Politique par service en DSL	[ADS09b][ADS09a] [LKB ⁺ 14]
Endormissement de lien		Politique de réveil	[CRN ⁺ 10][HARPFVLG11] [GS07b][GS07a] [eee10][GGS04]
Relais de proximité		Cas du réseau optique	[LKB ⁺ 14] [BKSP12]
		Application Transport	[JCN08][CGNG04] [SGRO ⁺ 08][GCN05]
Topologie active	Stratégie globale	Ingénierie de trafic verte dans IP	[LSC15][RRCRS14] [ZYLZ10][ACG13] [GS09]
		Ingénierie de trafic verte dans MPLS	[RRCRS15][FWM ⁺ 14] [ELCM13][CCHW11]
		Ingénierie de trafic verte dans SDN	[RRCRS15][CGM14] [WJG ⁺ 14][GMP14] [CGL14]
		Cas des data centers	[SLX10][AFLV08] [HSM ⁺ 10][LLW ⁺ 12]
		Approche distribuée	[CEL ⁺ 10] [CELP11]
		Approche centralisée	[CCPM12] [HJP15]

du réseau, en agissant sur les flux de trafic que doivent traiter les équipements. Si l'on empêche le trafic de circuler dans une zone, il est alors possible d'économiser de l'énergie dans cette zone. Les actions permettant d'agir sur la circulation du trafic s'appuient sur une connaissance du trafic, qui est supposée connue, lorsque l'on agit par planification du réseau, ou découverte lorsque l'on agit par

le plan de contrôle du réseau, c'est-à-dire le routage.

Dans cette thèse nous avons travaillé sur une stratégie locale et sur une stratégie globale avec pour la stratégie globale le souci de faciliter le déploiement grâce à une approche de commande logicielle du réseau, le Software Defined Network.

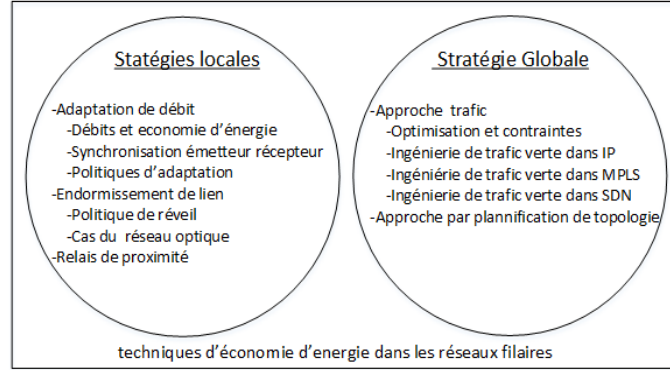


FIGURE 1.1 – Approche par stratégie globale versus stratégie locale des techniques d'économie d'énergie dans les réseaux filaires

Nous précisons à la suite le fonctionnement des quatre grandes approches relevées.

2 Stratégies locales d'économie d'énergie

2.1 Adaptation de débit (Adaptation Link Rate)

Avec cette technique une économie d'énergie est obtenue en réduisant le débit de transmission de la liaison (Figure 1.2). La gestion dynamique de la puissance modifie la tension des circuits lorsque la fréquence d'horloge (f) diminue entraînant une baisse de la consommation [CSB92]. Pour le processeur Complementary metal-oxide-semiconductor (CMOS) le modèle de puissance est approché par l'équation ci-dessous.

$$P = a \ C \ V_{dd}^2 \ f \quad (1.1)$$

Où P est la puissance dissipée, V_{dd} , la tension, a est le facteur d'activité, C est la capacité

équivalente et f est la fréquence de fonctionnement du processeur.

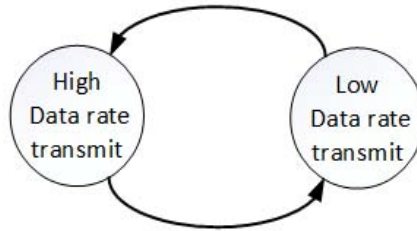


FIGURE 1.2 – Principe de l'adaptation de débit

L'approche par adaptation de débit qui est parfaitement adaptée à l'économie d'énergie d'un ordinateur doit en ce qui concerne le réseau prendre en compte l'aspect normalisation inhérent au domaine que ce soit au niveau des débits disponibles ou bien des échanges standardisés entre émetteur et récepteur.

2.1.1 Débits et économie d'énergie

Dans les réseaux, les débits de transferts sont normalisés et évoluent au cours du temps. Ainsi un transfert par une technologie Ethernet peut utiliser les gammes de 10 Mbps, 100 Mbps, 1 Gbps, 10 Gbps avec des horloges respectivement de fréquence 2.5 MHz, 25 MHz et 125 MHz. La fonction d'autonégociation des équipements Ethernet que nous présenterons dans le chapitre suivant, permet aux équipements de se mettre d'accord sur un des débits disponibles dans le standard. Le choix d'un débit moindre permet de diminuer la consommation d'énergie, ce que nous avons vérifié expérimentalement.

Dans [PDJC13], nous mesurons la consommation d'énergie sur des équipements DLink Switch (DGS1100-24), avec des débits de données allant de 10 Mbps à 1 Gbps. Pour mesurer les niveaux de puissance, nous utilisons un analyseur de puissance Rohde & Schwarz le HM8115-2 [HM8]. Les valeurs de consommation d'énergie que nous avons obtenues en configurant les ports d'un équipement à des débits différents sont présentées dans la figure 1.3. Les ports sont configurés, ils transmettent en permanence des symboles indiquant une liaison vide (c'est le fonctionnement de la transmission Ethernet).

Notons que sur cette figure nous trouvons également des valeurs de puissance avec l'application d'une stratégie locale d'économie d'énergie (l'endormissement), notée 1000E dans la légende de la figure 1.3. L'économie réalisée avec cette stratégie permet d'avoir la même puissance que celle résultant du choix d'un débit à 10 M.

Nous en concluons que transmettre lentement est moins coûteux en énergie que de transmettre rapidement et qu'il est possible de transmettre rapidement sans supplément de coût en utilisant un mécanisme d'économie d'énergie : l'endormissement.

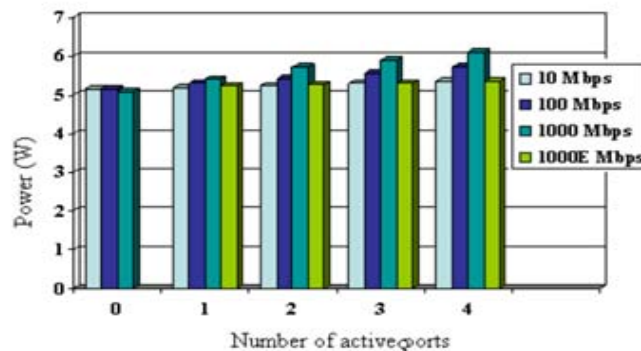


FIGURE 1.3 – Evaluation de la puissance consommée selon le débit en Ethernet [PDJC13]

2.1.2 Synchronisation émetteur récepteur

Un problème de l'adaptation de débit est la synchronisation qui doit être définie entre l'émetteur et le récepteur. L'équipement émetteur qui adapte son débit d'émission doit avertir l'équipement récepteur du lien. Pour favoriser la rapidité du déploiement d'équipements économes, il est souhaitable de réaliser l'échange en enrichissant des fonctions et standards déjà existants. Ceci peut être fait par une fonction d'autonégociation prévue dans la couche physique, ou un échange de trames du protocole d'accès [Media Access Control \(MAC\)](#). En ce dernier cas, le protocole standard doit avoir été conçu pour permettre de tels échanges, ce qui n'est pas le cas de la plupart des protocoles d'accès utilisés dans les réseaux filaires sur le segment d'accès ([Ethernet Passive Optical Networks \(EPON\)](#), [Gigabit-capable Passive Optical Networks \(GPON\)](#), [Data Over Cable Service Interface Specification](#)

(DOCSIS)). Un autre problème de la synchronisation par échange MAC est le temps nécessaire à l'adaptation. Il est plus important lorsqu'il est effectué à la couche liaison que lorsqu'il est effectué à la couche physique. [GCS06][ARSG06]

2.1.3 Politiques d'adaptation

Les travaux sur l'adaptation de débit mettent en exergue différentes mises en oeuvre pour décider du débit à utiliser. Nous relevons les mécanismes reposant sur des mesures d'utilisation de mémoire tampon et des mécanismes par type de service transféré.

2.1.3.1 Politique d'occupation de tampons Dans la littérature, le débit à utiliser est décidé à partir de la mesure de plusieurs paramètres : le taux d'occupation de tampon mémoire [GCS06][ARSG06], le taux d'occupation allié à la mesure de l'utilisation de la liaison [ARSG06][GCN05][NPI⁺08], ou encore des mesures de délai [GCNS08].

(i) **Politique à deux seuils** : ALR **Dual-Threshold Policies (DTP)** [ARSG06] est une politique à deux seuils qui sont associés au taux de remplissage d'un buffer stockant les paquets à transmettre. Lorsque l'état d'occupation du buffer est supérieur au seuil maximal, le débit de la liaison est augmenté, il est diminué lorsque le taux d'occupation est inférieur au seuil minimal. Les auteurs ont observé que si les valeurs de seuil bas et haut de tampon sont ajustées à 50% et 80% respectivement, alors le lien peut être exploité au débit le plus bas pratiquement à 100% avec un retard de paquets négligeable.

(ii) **Politique à seuils avec mesure de la liaison** : Gunaratna et al. [GCN05] améliorent la précision de l'adaptation en adjoignant aux mesures mémoire qui sont fortement influencées par la taille du tampon, la mesure de l'occupation de la ligne. Ceci permet de diminuer les oscillations de changements d'états.

(iii) **Politique à seuils avec délai** : Dans [GCNS08] il est proposé de lisser davantage la politique d'adaptation en la conditionnant par des intervalles de temps prédéfinis. Si la liaison est

en débit haut depuis peu de temps même si le buffer est quasi vide le débit ne sera pas changé.

2.1.3.2 Politique par service en DSL Dans le réseau d'accès filaire cuivre en technologie DSL ([Digital Subscriber Line \(DSL\)](#)), qui est la technologie actuellement dominante sur ce segment de l'architecture large bande, des mécanismes d'adaptation de débit sont associés à des services. Les services sont ceux offerts par l'opérateur d'accès pour une offre triple play : voix, données, image. L'économie d'énergie est réalisée par l'utilisation d'un mode de consommation basse puissance qui a été standardisé pour les accès en [Asymmetric Digital Subscriber Line 2 \(ADSL2\)](#) et est en cours de standardisation pour les accès [Very High Speed Digital Subscriber Line 2 \(VDSL2\)](#) et pour les accès optique en [GPON](#).

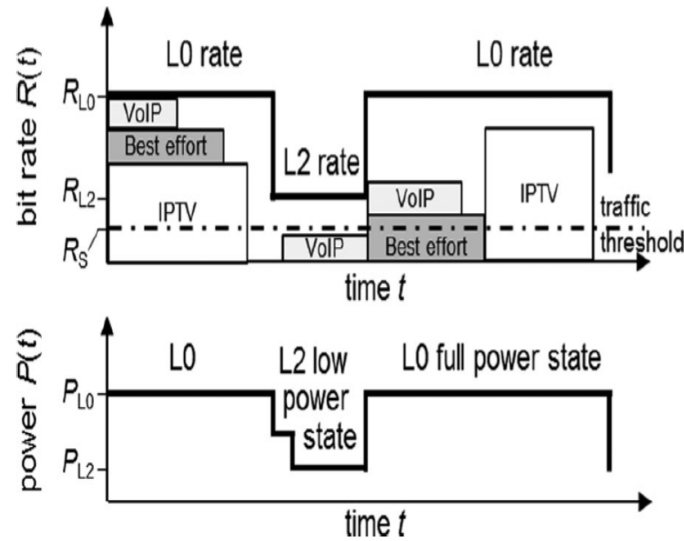
Le principe de fonctionnement [[ADS09b](#)][[ADS09a](#)] illustré en figure 1.4 est le suivant :

Deux niveaux de puissance sont utilisés : P_{L0} et P_{L2} correspondant à deux débits de transfert R_{L0} et R_{L2} qui peuvent être calculés selon l'offre de service de l'opérateur. Les niveaux sont utilisés de façon locale et autonome sur la ligne et négociés entre le modem [ADSL](#) et son [Digital Subscriber Line Access Multiplexer \(DSLAM\)](#) (figure 2). Le trafic transmis est comparé à des seuils préfixés. Si le trafic descend sous le seuil noté R_s en figure 1.4, le débit qui est utilisé est R_{L2} . Si le trafic est supérieur à R_s le débit utilisé est $L0$.

Sur l'illustration de la figure 1.4, le débit faible est utilisé (R_{L2}) pour un service de voix sur IP qui correspond à un transfert de données inférieur au seuil minimum. Le débit plus élevé (R_{L0}) est proposé pour le service de transmission de programme télévisuel par internet (IP TV). Le débit R_{L2} est également employé en cas d'une utilisation simultanée des services.

L'économie d'énergie qui est obtenue par cette méthode est fonction des profils de trafics des utilisateurs, les auteurs de [[LKB⁺14](#)] se réfèrent à des statistiques de trafic et pour celles-ci indiquent une économie d'énergie de 23% .

Conclusion : La fonction d'adaptation de débit s'applique à toutes les technologies de transmission et sur les différents segments du réseau. Pour un segment d'accès cuivre en transmission [DSL](#),

FIGURE 1.4 – Principe du fonctionnement de l'ADSL avec économie d'énergie [LKB⁺14]

la standardisation a opté pour cette technique. Par contre pour le segment Ethernet, qui fait l'objet de nos travaux ce n'est pas le cas, cette technique étant jugée plus coûteuse que la technique d'endormissement que nous présentons à la suite et que nous détaillerons plus précisément dans le chapitre suivant.

2.2 Endormissement de lien

Lorsqu'un équipement de transmission n'a pas de données à transmettre il consomme de l'énergie. Comme pour les ordinateurs, un mécanisme de mise en sommeil permet de diminuer la consommation. Ainsi par exemple le standard Ethernet préconise un état de veille consommant 10% de la puissance consommée en état actif.

2.2.1 Politique de réveil

Le principe du mécanisme de sommeil est très simple. La mise en sommeil s'effectue lorsqu'il n'y a pas de paquet à transmettre, et s'arrête lorsqu'il y a des paquets à transmettre. Cependant sa mise en oeuvre est variée. Les options se déclinent selon plusieurs caractéristiques relatives au nombre

nombre d'états, à la politique d'endormissement et de réveil. (Figure 1.5)

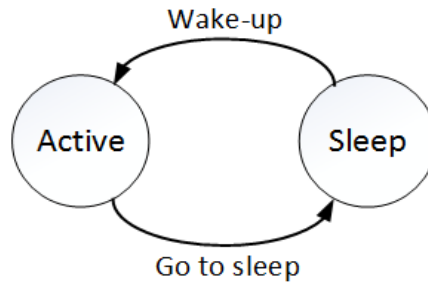


FIGURE 1.5 – Approche simple à deux états de l'endormissement de lien.

Le passage à l'état actif n'est pas instantané [GGS04] [AK08]. Selon les options envisagées le paquet peut être stocké, le temps que l'état actif soit effectif.

L'objectif des travaux présentés dans la littérature est de maximiser le temps de sommeil. Par exemple, les paquets peuvent être stockés, comme pour le mécanisme d'adaptation de débit, ce qui permet d'agréger l'émission en transmettant plusieurs paquets en suivant et augmente le temps de sommeil.

De nombreux travaux ont été menés sur la politique de réveil [eee10] [GS03] [GGS04] [GS07a] [HARPFVLG11] [AK08] [CRN⁺10] [GS07b]. Plus précisément les politiques décrites en [GGS04] considèrent trois fonctionnements pour un équipement réseau :

- a) une interface dort durant un temps donné, T . Durant cette période tous les paquets devant être traités par l'interface seront perdus.
- b) une interface est en sommeil, l'arrivée d'un paquet la réveille mais elle n'est pas active immédiatement et donc le paquet qui l'a réveillé est perdu. Les paquets suivants seront traités.
- c) une interface est en sommeil, l'arrivée d'un paquet la réveille et en attendant que l'interface soit prête à traiter le paquet celui-ci est mis en attente. La mise en attente du paquet se fait

dans un tampon géré par un matériel spécifique. L'interface en sommeil se réveille cycliquement pour vérifier si il y a quelque chose à transmettre. Le standard d'endormissement de lien Ethernet s'appuie sur cette mise en oeuvre.

Des politique de reveil plus évoluées ont été proposées qui rappellent les politiques d'occupation des tampons présentées précédemment.

Katz et Ananthanarayanan dans leur étude de verdissement des switch [AK08] proposent de stocker les paquets à émettre, lorsque l'équipement est en sommeil et d'introduire une interface gérant un buffer pour stocker un paquet en réception. Plutôt que de se réveiller de façon cyclique, les auteurs proposent un réveil à l'arrivée d'un paquet. Ainsi lorsqu'il n'y a pas de trafic, le switch peut rester endormi. Pour améliorer cette politique ils élaborent un mécanisme de réveil prédictif. Ce mécanisme observe le nombre de paquets qui arrivent durant une fenêtre de temps et, considérant que l'indication de trafic obtenue est valable pour la prochaine fenêtre de temps, décide de la mise en sommeil de l'équipement. Si le trafic est inférieur à un seuil, les paquets sont mémorisés pour être transmis au prochain réveil.

Gupta et Singh dans leurs travaux pionniers [GS03] avaient également utilisé un mécanisme basé sur le trafic pour déterminer le déclenchement de l'endormissement. Ils mesuraient le temps d'arrivée inter paquets et basculaient en état de sommeil au-delà d'un seuil. Pour diminuer les pertes de paquets les auteurs ont proposé d'émettre un paquet de réveil aux voisins pour permettre à ceux-ci d'avoir le temps de faire la transition entre l'état de sommeil et l'état actif.

2.2.2 Cas du réseau optique

L'architecture du réseau d'accès optique (figure 2) est composée d'un élément réseau **Optical Line Terminal (OLT)** qui donne l'accès réseau aux utilisateurs au travers d'équipements modem nommés **Optical Network Unit (ONU)**. Les travaux en économie d'énergie concernent les équipements ONU et l'équipement réseau l'OLT.

Pour les équipements ONU, un moyen de diminuer la consommation a été proposé par Alcatel [Lucent] reposant sur le protocole Bipop [SCvV⁺12].

Grâce à un mécanisme spécifique d'entrelacement des informations transmises sur le réseau par l'OLT, l'équipement ONU qui de par la topologie en bus diffusé du réseau optique passif reçoit toutes les informations, est capable de détecter très rapidement celles qui le concernent et de détruire les autres. Le protocole définit un format de trame descendant (émise par l'OLT à destination des ONU) qui reste de 125 μ s mais où les bits d'entête et de données des ONT sont régulièrement répartis dans la trame, au lieu d'être positionnés à des endroits variables. L'ONU récupère plus rapidement ses données et détruit les autres. Il économise des traitements inutiles.

Pour économiser l'énergie au niveau de l'OLT deux méthodes sont étudiées par la standardisation : l'endormissement et la mise sous tension de délestage

1) La méthode d'endormissement est proposée par la standardisation ITU-T G.987.3 [XG-14] sur le réseau optique à 10 Gigabits/10-Gigabit-capable Passive Optical Network (X-GPON). L'objectif est de pouvoir éteindre les circuits relatifs aux fonctions optiques, MAC, et stockage lorsqu'il n'y a pas de trafic à transmettre.

Deux états d'endormissement sont spécifiés dans le standard.

- Le premier désactive sur l'équipement utilisateur les fonctions de transmission mais laisse actives celles de réception, l'équipement ONU est alors capable de recevoir un message de l'OLT pour se réveiller, il est en état de sommeil léger (dozen state). L'illustration du fonctionnement est donné sur la figure 1.6.
- Le deuxième mode désactive les fonctions d'émission et de réception. Comme les informations sont physiquement diffusées à tous les ONU, l'arrivée d'un signal ne peut être le déclenchement de la période d'activité. C'est pourquoi l'équipement se réveille de façon cyclique. Il est en état d'endormissement cyclique (Cycle sleep). La synchronisation entre l'émetteur et le récepteur se fait par un échange d'information de gestion l'OAM (Operation And Maintenance (OAM)).

2) La méthode de désactivation (ou shedding, pour puissance de délestage) repose sur l'arrêt des services. La recommandation G.984.4 contient déjà cette capacité d'arrêt de services inutiles en cas



Pour arrêter les services, l'OLT mesure l'activité du lien et déclenche l'arrêt de l'interface inutilisée.

Les fonctions du coeur relatives à une interface qui est éteinte sont alors également arrêtées. Les ports virtuels qui sont associés aux types de trafic reçus par l'utilisateur au travers de l'OLT peuvent être déconnectés. Ces ports restent provisionnés et peuvent être rétablis à la demande de l'utilisatrice.

[LKB⁺14] présente en se référant à des statistiques de trafic sur un réseau résidentiel, une comparaison des méthodes d'endormissement léger, endormissement cyclique ainsi que la méthode par désactivation délestage qui fait apparaître un gain en faveur de l'endormissement cyclique.

L'endormissement cyclique est modélisé mathématiquement en [BKSP12] [KZY⁺14]. Ils en déduisent une durée de sommeil optimale. Au-delà de cette période, il n'y a pas de gain d'énergie et le délai est augmenté.

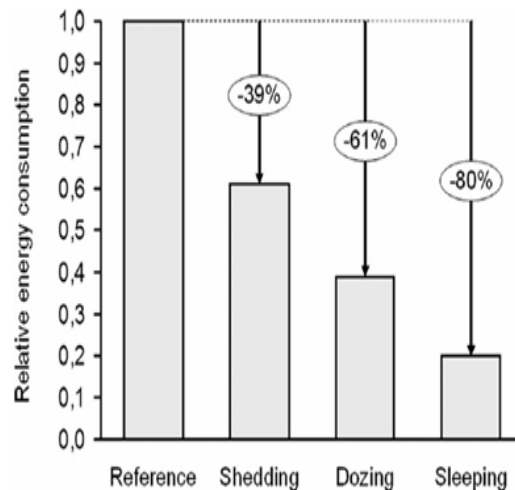


FIGURE 1.7 – Comparaison des méthodes d'endormissement optique [LKB⁺14]

Nous venons de décrire le fonctionnement du réseau optique économe XPON normalisé par l'ITU(International Telecommunication Union (ITU)), nous verrons dans le chapitre suivant le fonctionnement normalisé par l'IEEE qui s'applique au réseau optique EPON qui quoique similaire au réseau GPON ne repose pas sur un contrôle du réseau par OAM. L'IEEE a choisi une méthode d'endormissement qui ne fait pas appel à un échange protocolaire OAM mais à une fonction déjà existante dans les versions Ethernet cuivre 100Mbps l'auto négociation.

2.3 Relais de proximité

Le principe de la méthode est de recourir à un équipement intermédiaire, le relais de proximité, pour travailler à la place de un ou plusieurs équipements. L'équipement intermédiaire est typiquement un switch/routeur du réseau d'accès qui va permettre à ses clients de dormir.

Ce fonctionnement est de plus haut niveau que ceux présentés précédemment. Il s'agit de travailler à la place d'une entité pour la laisser dormir, en répondant à un protocole à sa place. Par exemple, la réponse à une requête de recherche d'identité par la protocole ARP ([Address Resolution Protocol \(ARP\)](#)) est faite par l'élément de proximité. Cette solution peut être mise en oeuvre dans les cartes réseaux comme proposé par Subha Natarajan et al. [[SGRO⁺08](#)].

Des exemples de fonctions que peut assurer un relais de proximité sont présentés dans les travaux [[JCN08](#)] [[WMG10](#)] :

- **Network Connectivity Proxy (NCP)** : Les auteurs dans ce papier se préoccupent de la gestion des connexions [Transmission Control Protocol \(TCP\)](#). Lorsqu'une connexion [TCP](#) est ouverte une entité qui ne s'en sert pas doit la maintenir ouverte par l'envoi de messages Keep alive. Pour lui permettre de dormir le relais de proximité se charge d'envoyer les message à sa place.

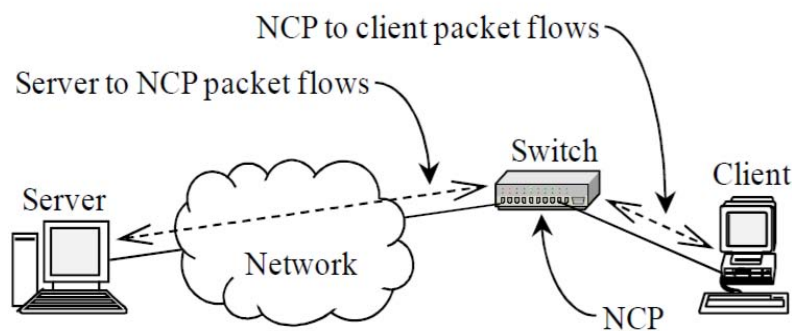


FIGURE 1.8 – Illustration de la méthode de sommeil avec relais de proximité.

- **Proxy NICs** : Wang et al. [WMG10] s'intéressent à l'application de messagerie instantanée Yahoo Messenger. Le but est de dormir pendant les phases d'inactivité. Sachant que les périodes d'inactivité de l'application, sont aléatoires, ils formulent le problème de décision avec un processus stochastique et proposent un algorithme adaptatif pour configurer les valeurs d'endormissement.

3 Stratégie Globale d'économie d'énergie

Nous venons de présenter des travaux portant sur certains éléments du réseau, nous considérons à présent le réseau dans son ensemble et, comment sa topologie peut être définie pour économiser de l'énergie. Les travaux relatifs à la définition d'une topologie active, c'est-à-dire constituée d'équipements en activité nécessitent une stratégie globale du réseau avec une coordination entre ces éléments. Il s'agit d'agir sur le trafic pour lui faire prendre un chemin passant par des équipements qui consommeront de l'énergie, alors que les chemins sur lesquels le trafic ne passera pas seront source d'économie d'énergie. Ces économies seront naturellement fonction du modèle de consommation de la technologie utilisée sur les chemins.

Les travaux considèrent dans leur grande majorité un modèle de consommation indépendant de la technologie sous-jacente de transmission.

Quelle que soit la technologie nous pouvons considérer des économies relevant d'une mise en sommeil et celles liées à un arrêt de l'équipement. La méthode d'endormissement permet de traiter du trafic et de s'endormir entre plusieurs informations alors que l'arrêt permet de ne pas traiter le trafic sauf en cas de problème. La différence entre les deux réside essentiellement dans l'échelle de temps, temps plus court pour l'endormissement que pour l'arrêt. L'une relève du contrôle du réseau l'autre davantage de son administration.

C'est pourquoi nous séparons deux approches :

- celles relevant de l'ingénierie de trafic gérées par le plan contrôle réseau, elles s'appuient sur

le routage, et

- celles relevant de la conception de la topologie mise en oeuvre par plan de gestion qui planifie un réseau économe en énergie.

Les premières approches utilisent des méthodes d'optimisation avec contraintes, les secondes s'appuient sur des méthodes de graphes.

3.1 Approche trafic

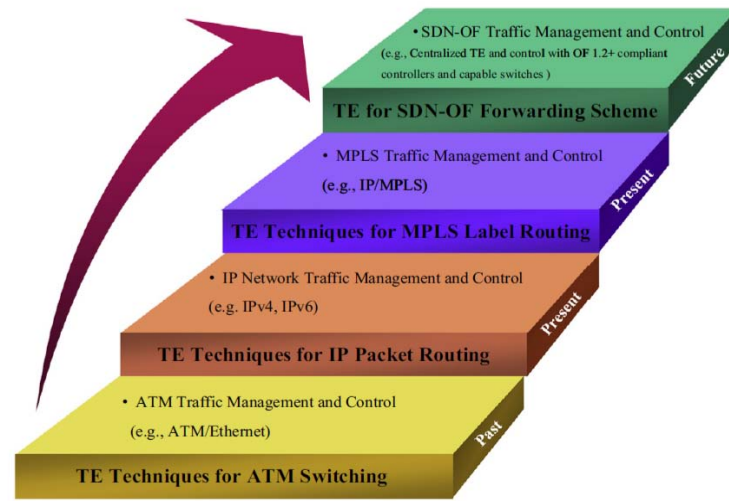
L'approche consiste à diriger le trafic sur un chemin choisi pour économiser de l'énergie par rapport à un chemin calculé sur la seule performance traditionnellement exprimée en nombre de noeuds.

Les moyens d'y parvenir sont des moyens d'ingénierie de trafic tels que décrits dans le RFC 3272 de 2002 [TE02]. L'ingénierie du trafic consiste à adapter l'acheminement du trafic aux conditions du réseau, avec un objectif de performance pour l'utilisatrice et d'utilisation efficace des ressources pour l'opérateur du réseau. Les travaux dans ce domaine (Figure 1.9) ont été initiés dans le cadre de la technologie ATM (*Asynchronous Transfer Mode (ATM)*), avec déjà le souci de la performance de consommation étudiée en [PN14] puis repris dans le cadre des architectures IP avec des travaux considérant des trafics sous forme paquets [GS09][ZYLZ10][ACG13][LSC15][RRCRS14], et des travaux considérant des trafics sous forme de flux avec MPLS [CCHW11][FWM⁺14][ELCM13][RRCSR15].

Les travaux les plus récents sont relatifs à l'utilisation de contrôleurs centralisés pour gérer le routage dans les approches de réseau par SDN [WJG⁺14][GMP14][CGM14][RRCS15]. Nous décrirons plus précisément cette technologie dans le chapitre 4 de la thèse.

3.1.1 Optimisation et contraintes

L'ingénierie de trafic s'est initialement développée dans le but d'éviter les congestions dans le réseau opérateur et de respecter la qualité de service pour les utilisateurs. Il s'agit de définir le routage des informations dans le réseau avec l'objectif de minimiser le délai en respectant une contrainte sur

FIGURE 1.9 – Historique de l'ingénierie [ALW⁺14].

la charge des liaisons afin d'éviter de perdre des données. Par la suite, de nombreux travaux ont été réalisés pour adapter le routage à des objectifs variés, les références de tels travaux seront trouvées en [WHPH08]. Généralement la problème de diminution de l'énergie consommée se traduit par un problème d'optimisation similaire qui considère un objectif de minimisation de l'énergie.

Gupta et al dans leurs travaux pionniers en 2003 [GS03] identifient le problème d'économie d'énergie et émettent l'idée de mettre en sommeil des équipements. Une approche de mise en oeuvre de cette idée qui repose sur un contrôleur centralisé qui collecte les données de topologie et de trafic, est proposée en [GS09]. L'approche peut être mise en oeuvre avec des protocoles IP/OSPF et MPLS. Le contrôleur calcule les meilleurs chemins en maximisant le nombre de liens qui peuvent être mis en sommeil avec une contrainte d'utilisation de liens qui permet d'éviter les congestions, elle est par défaut évaluée à 50%. Le problème est formulé en MCF (**M**ulti **C**ommodity **F**low (**MCF**)) avec un objectif de minimisation de l'énergie globale sur le réseau par mise en sommeil des éléments qui n'acheminent pas d'information. Lorsque tous les ports d'une carte sont en sommeil, la carte peut être mise en sommeil. Le coût en énergie est celui de la carte d'interface plus celui des ports actifs sur chaque carte, il est indépendant de la charge de trafic par lien. Les demandes de transfert d'information sont réparties en partage de charge, sur un ensemble de plus courts chemins (une limite sur la

taille de chemins permet de respecter une contrainte de délai). La résolution est obtenue avec l'outil CPLEX, sur des topologies réelles (comme Géant ou Abilene) le temps étant estimé trop important la résolution est stoppée après 300s. Pour faire face à cette durée importante de calcul, le routage s'effectue sur des chemins précalculés.

L'étude d'heuristiques de résolution pour le problème d'optimisation NP difficile est abordé par Chiaraviglio et al [CMN12]. Ils modélisent le problème de la minimisation de l'énergie avec résolution en programmation linéaire en nombre entier et proposent des heuristiques de résolution, comme une réduction de l'espace de calcul qui considère que les cartes des sources et des destinations de trafic ne peuvent pas être éteintes.

En 2015, Lin et al. [LSC15] étudient des contraintes de fiabilité. Ils proposent **Reliable Energy-Aware-Routing (R-EAR)** qui introduit deux contraintes de fiabilité celle du terminal et celle du chemin : (i) Terminal Reliability (TR), (ii) Route Reliability (RR). TR_d est la probabilité qu'il existe au moins un chemin opérationnel entre la source et la destination pour la demande d . RR_d , est la probabilité d'avoir une route opérationnelle lorsqu'il y a une rupture de lien. Les auteurs utilisent un solveur en programmation linéaire mixte pour résoudre le problème et rajoutent deux contraintes au problème MCF à savoir :

$$TR_d \geq TR_d^* \quad , \quad RR_d \geq RR_d^* \quad (i, j) \in E \quad (1.2)$$

Soient TR_d^* et RR_d^* les fiabilités du réseau original. Les contraintes exprimées en (1.2) indiquent que la fiabilité du réseau au niveau terminal et au niveau route qui sera obtenue après avoir enlevé certains liens doit être supérieure à un certain multiple de la fiabilité d'origine avec $0 < \alpha < 1$.

3.1.2 Ingénierie de trafic verte dans IP

Dans [SLX⁺12] Shen et al s'intéressent à la réalisation d'un routage efficace en énergie ROD (**Routing On Demand (ROD)**) avec une optimisation de l'énergie et une répartition du trafic qui minimise la charge des liens alors que dans de nombreux travaux il s'agit de considérer une contrainte de charge maximale sur le lien. Pour ce faire ils introduisent un facteur de verdissement dans la

fonction objectif. Ils s'appuient sur un contrôleur central qui calcule périodiquement les poids et configure les routeurs.

Shen et al. prouvent qu'il existe toujours un ensemble de poids pour chaque objectif et que les solutions, c'est-à-dire les routes, peuvent être converties en plus court chemin réalisables avec le protocole de routage à état de liaison standard de l'Internet, [OSPF](#). Les poids des liens sont calculés en programmation non linéaire.

Dans [\[ACG13\]](#), les auteurs proposent de changer régulièrement les poids [OSPF](#) en s'appuyant sur des matrices de trafic exprimées à plusieurs périodes de la journée (figure 1.10). Ils minimisent la consommation d'énergie et la congestion du réseau en respectant des contraintes de qualité de service pour acheminer le trafic. Les poids [OSPF](#) sont calculés hors ligne par une heuristique de résolution du problème MILP ([Mixte Integer Linear Programming \(MILP\)](#)).

Les auteurs divisent la journée en trois périodes caractérisées par un niveau quasi stable de trafic.

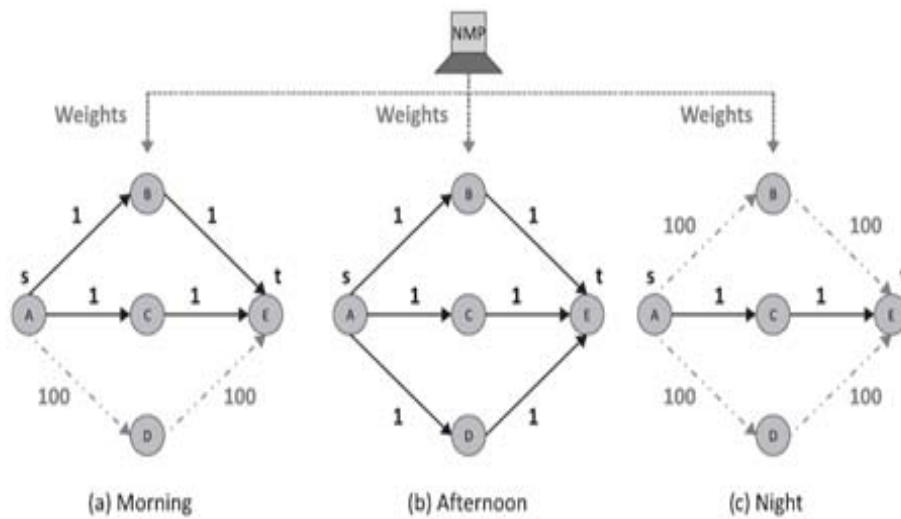


FIGURE 1.10 – Exemple de calcul de poids pour un routage à énergie selon 3 périodes de la journée : a) matin, b) après-midi c) soir [\[ACG13\]](#).

Ruiz-Rivera et al. [\[RRCRS14\]](#) s'intéressent aux effets de l'optimisation sur le fonctionnement

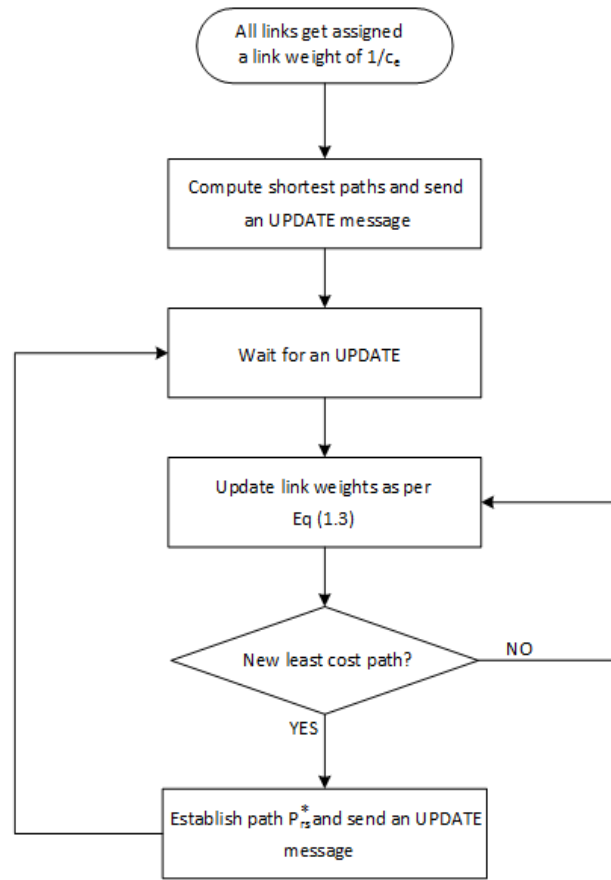


FIGURE 1.11 – Fonctionnement du routage HotPLUZ [RRCRS14]

de l'internet, et comment le changement de routage interne affecte le protocole inter systèmes autonome BGP ([Border Gateway Protocol \(BGP\)](#)). En reprenant les propositions de calculs de poids de [ACG13], ils évaluent à 29% de plus, le nombre de changements de routes par un routage économe en énergie comparé à un routage classique, ce qui augmente d'autant les possibilités de perte de paquet lors des calculs de table de relayage, les congestions sur les chemins entre systèmes autonomes dues au glissement du trafic. Leur proposition Hotpluz inspirée de celle de [TSGR08] pour augmenter la stabilité des routes, calcule les poids de façon à diriger le trafic des liens faiblement chargés vers les liens fortement chargés. Le calcul est donné en Equation 1.3. La figure 1.11 illustre le fonctionnement du routage. A l'initialisation les routeurs internes calculent le chemin le plus court en nombre de

saut jusqu'à leur routeur de sortie. Chaque routeur interne prévient les autres routeurs de son calcul (message UPDATE), les routeurs utilisent ces données pour mettre à jour leur calcul. Le processus s'arrête lorsqu'il n'y a plus de changement, les routeurs ont convergé sur le plus court chemin. A la fin du processus les liens sans trafic peuvent être désactivés par l'opérateur.

$$e = \begin{cases} \frac{1}{1-(C_e-L_e)} & , \text{if } 0 < L_e \leq MLU \\ & , \text{if } L_e > MLU \end{cases} \quad (1.3)$$

3.1.3 Ingénierie de trafic verte dans MPLS

L'acheminement des informations par [MPLS](#), c'est-à-dire dans un mode connecté, rend plus aisé les opérations d'ingénierie de trafic de mesure et contrôle de trafic, et grâce à l'établissement de chemins en mode contraint via un routage explicite par la source, il est possible de diriger le trafic sur des chemins. C'est une solution attrayante pour mettre en place des politiques vertes comme en attestent les travaux [[CCHW11](#)][[FWM⁺14](#)][[ELCM13](#)][[RRC SR15](#)]. Par rapport aux travaux précédents, une dimension supplémentaire est ajoutée par l'établissement des connexions [MPLS](#). Etablir des connexions prend du temps et des ressources il y a donc intérêt à ne pas ouvrir fermer des connexions trop fréquemment.

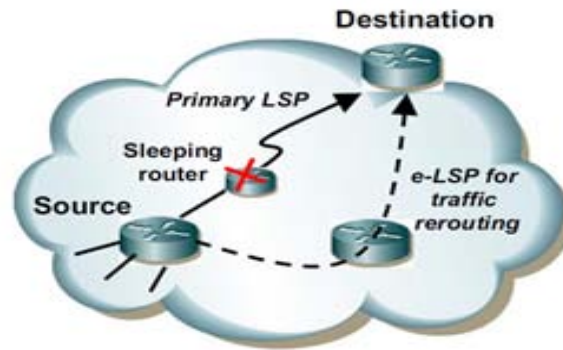


FIGURE 1.12 – Reroutage de traffic sur LSP économe en énergie (e-LSP) [[CCHW11](#)].

Green MPLS Traffic Engineering proposé dans [[CCHW11](#)] répartit le trafic pendant les périodes

creuses vers des liens chargés, tout en évitant de créer trop de connexions MPLS (LSP) afin éviter des temps d'établissements de connexion pénalisants. L'approche [FWM⁺14] utilise les connexions de secours installées dans un réseau MPLS pour se protéger des pannes. En cas de panne d'un lien, le trafic est reporté sur le chemin de secours. Avec cette approche il n'y a pas besoin de gérer des établissements de connexions, elles sont déjà installées. L'optimisation consiste à répartir au mieux le trafic pour économiser l'énergie en ne dépassant pas la capacité des chemins de protection qui en cas de panne retrouvent leur fonction première de protection.

[RRCSSR15] s'intéresse plus précisément à l'établissement des connexions. Il rajoute aux critères d'acceptation des connexions classiques que sont le débit et les paramètres de qualité de service (délai, perte) un paramètre d'énergie. Avec un taux d'acceptation de 90% ils obtiennent sur des topologies Abilene et AT&T un taux d'endormissement de liens de 20% en dehors des périodes de pointe.

3.1.4 Ingénierie de trafic verte dans Software Defined Network

Le réseau SDN que nous présenterons plus précisément dans le chapitre 4, repose sur des décisions de contrôle centralisées en un point, le contrôleur. L'architecture du réseau est composée d'éléments de relayage, les switches, qui communiquent avec un contrôleur. Celui-ci est programmable, et met en oeuvre les souhaits de l'administrateur du réseau qui décide de la politique à appliquer. L'ingénierie de trafic est réalisée par le contrôleur qui indique aux éléments de relayage, sur quel port doit être transmis une information entrante. La décision est mémorisée pour un temps programmable par Timer dans la table de relayage de l'équipement, ce qui évite à ce dernier de devoir demander à son contrôleur ce qu'il doit faire à chaque information entrante. La décision de routage n'est plus effectuée sur chaque équipement comme dans une architecture traditionnelle de relayage IP ou Ethernet, elle est centralisée sur le contrôleur.

Les commutateurs SDN peuvent être programmés via le contrôleur de façon proactive, dans une phase d'initialisation, préalable à l'échange de trafic, et/ou programmés de façon dynamique pour répartir de façon optimale les ressources du réseau.

Wang et al. [WJG⁺14] étudient le problème d'optimisation de l'énergie en SDN par reroutage du trafic, de façon similaire à l'approche de Green-TE en y ajoutant une diminution du temps de calcul

de la solution optimale par heuristique.

[CGM14] s'intéresse au réseau SDN avec MPLS. Les travaux supposent qu'un ensemble de chemins MPLS sont pré calculés et installés. Le contrôleur effectue le contrôle d'admission. Les auteurs proposent un algorithme de placement des flux sur des chemins actifs qui agrégé le trafic afin de mettre en sommeil les équipements, ainsi qu'un algorithme pour répartir la charge des chemins congestionnés sur plusieurs flux.

Dans [GMP14], le problème d'optimisation prend en compte une contrainte sur la taille des tables des équipements de relayage. Réalisée en TCAM (Ternary Content Adressable Memory (TCAM)), celle-ci est actuellement de l'ordre de 1000 entrées. Les auteurs montrent que la limitation sur la taille des règles mémorisées par les éléments SDN peut affecter le routage. Une illustration en est donnée Figure 1.13, y sont représentés : la topologie du réseau avec la capacité des liens, les résultats de routage obtenus avec un routage à économie d'énergie sans contrainte sur le nombre de règles et, avec contrainte.

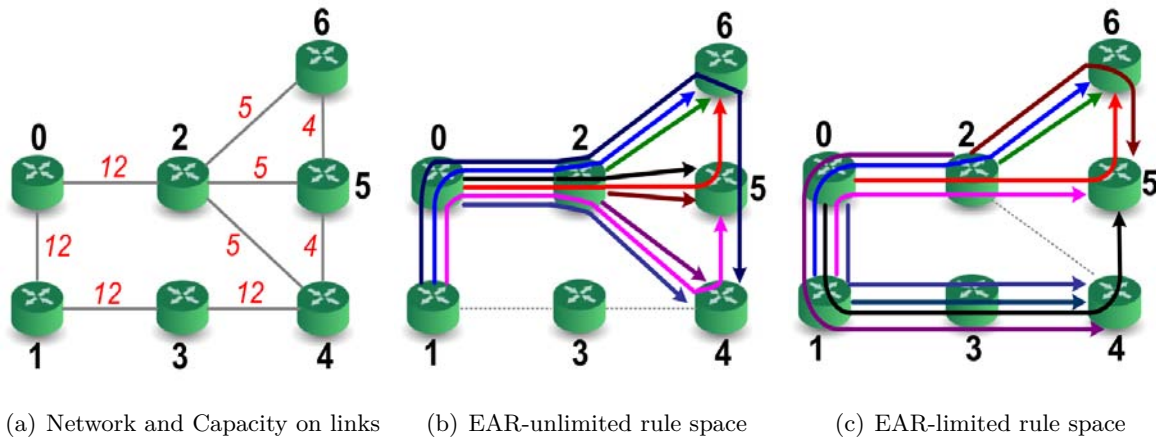


FIGURE 1.13 – Influence du nombre de règles SDN sur le routage à économie d'énergie [GMP14].

En 2015, les auteurs [RRCS15] s'intéressent à un réseau multi contrôleurs, qui fonctionne avec un contrôle distribué entre les contrôleurs [SS13], et étudient la répartition des switches et des contrôleurs pour économiser l'énergie. La modélisation a pour but de mettre en sommeil le maximum de lignes, mais cette fois en s'assurant que les switches ont tous une route avec un contrôleur qui a un délai

limité et que les contrôleurs ont une charge limitée afin qu'ils travaillent le plus rapidement possible.

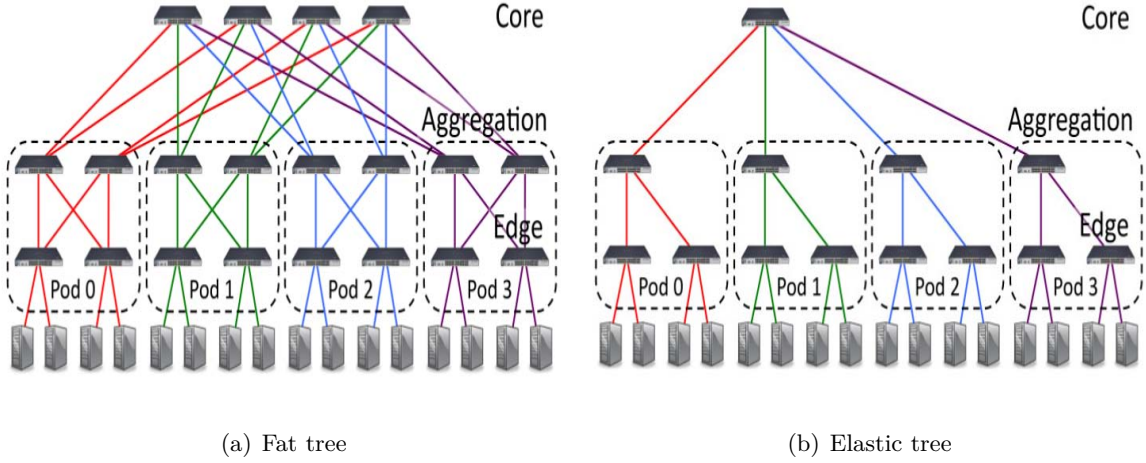
3.1.5 Cas des data centers

Dans le domaine des réseaux de Data center (Data Center Network), les travaux en économie d'énergie peuvent être séparés entre ceux portant sur le placement des machines virtuelles, et ceux sur le routage des flux. Parmi ces derniers nous retrouvons des approches similaires à celles vues précédemment, mais appliquées aux topologies particulières des [Data Center Network \(DCN\)](#). Les topologies réseaux des Data center doivent interconnecter un nombre très important de serveurs et offrir un haut niveau de disponibilité avec une redondance pour fiabiliser la communication des applications de haut niveau. La topologie doit être également facilement évolutive pour raccorder de nouveaux éléments. Pour répondre à ces besoins et améliorer l'architecture classique dite three tiers [[Cis07](#)], de nouvelles architectures ont été proposées [[AFLV08](#)][[SLX10](#)][[GLL⁺09](#)] et étudiées sous le point de vue consommation énergie [[HSM⁺10](#)][[LLW⁺12](#)].

Elastic tree de Brandon et all [[HSM⁺10](#)] est un système qui adapte dynamiquement la consommation de la topologie [DCN](#). A partir de statistiques sur l'utilisation du système, il calcule une solution optimisée de la topologie, en considérant que les flux ne sont pas répartis sur différents chemins, et proposent deux heuristiques de résolution adaptées à la spécificité de la topologie [DCN](#). L'une choisit le chemin le plus à gauche dans la topologie (figure [1.10](#)) ayant la capacité suffisante, l'autre s'appuie sur la symétrie de la topologie, en déterminant le nombre d'éléments nécessaires pour transmettre un trafic plutôt que de déterminer quel élément doit être utilisé, ce qui diminue la complexité du calcul. Une fois l'optimisation calculée le routage est mis en oeuvre par des switch [SDN](#). Une illustration de la topologie obtenue est donnée en Figure [1.14](#). D'autres heuristiques seront trouvées en [[WJZW12](#)][[LGYN13](#)].

3.2 Approche par planification de topologie

Les approches précédentes s'appuyaient sur la connaissance du trafic. Les travaux présentés utilisaient des matrices réelles provenant de mesures effectuées sur des réseaux tels que Abilene, Géant ou Orange, [[ZYLZ10](#)] [[CGM14](#)] [[ACG13](#)] ou bien sur des modélisations de trafic [[RRCRS14](#)] [[CCHW11](#)].

FIGURE 1.14 – Exemple de topologie en réseau DCN [HSM⁺10]

Nous présentons à présent des travaux sans connaissance de trafic, avec une vision centralisée puis distribuée.

3.2.1 Approche distribuée

En mode distribué chaque routeur détermine les liens qu'il peut désactiver, grâce à des échanges avec ses voisins. Les échanges sont généralement effectués avec [OSPF](#). Cianfrani et al [CEL⁺10] [CELP11] ont proposé une approche compatible avec le protocole [OSPF](#) qui se concentre sur le calcul de l'arbre des plus courts chemins effectués par les routeurs, pour déterminer leurs routes, indépendamment des trafics. Leur proposition vise à réutiliser les plus courts chemins de ses voisins pour laisser inutilisés des liens et les mettre en sommeil. Dans un réseau classique chaque routeur calcule son arbre, dans leur proposition, nommée [Energy Aware Routing \(EAR\)](#), seuls certains routeurs font ce calcul et, les autres routeurs importent ce résultat. Les routeurs importateurs utilisent l'arbre importé en calculant leur chemin jusqu'à leur exportateur qui est considéré comme la racine de l'arbre des plus courts chemins. Des travaux sur la répartition importateur/exportateur sont présentés en [CELP11].

Bianzino et al.s'appuient sur une méthode d'apprentissage par renforcement pour que les routeurs décident d'activer ou pas leurs liens. Les échanges [OSPF](#) permettent à un noeud de connaître l'état

du système, modélisé par 3 états (ligne inutilisée, ligne utilisée, ligne congestionnée) avec une fonction d'utilité sur l'énergie. La méthode n'a pas besoin de connaître le trafic avec sa source et sa destination, elle calcule la charge d'un lien, ce qui peut être fait à chaque routeur.

3.2.2 Approche centralisée

Les travaux [CCPM12] [CAP11] [HJP15] supposent une vue globale de la topologie qui peut être obtenue par le système de gestion du réseau ou encore les bases de routage des routeurs. Il s'agit de déterminer les éléments les moins importants de la topologie, de façon à pouvoir les enlever de la topologie pour obtenir une topologie active économe en énergie.

Les travaux de cette approche utilisent plusieurs outils des graphes. L'approche proposée par Cuomo et al [CAP11] repose sur une analyse des plus courts chemins pour déterminer les éléments lignes et nœuds qui y apparaissent le moins souvent, afin de pouvoir les désactiver. [CCPM12] et [HJP15] utilisent la théorie des graphes pour déterminer ces éléments, ils sont déterminés par des calculs de connectivité algébrique et de centralité.

4 Conclusion

Les travaux en transmission à économie d'énergie avec une stratégie locale que nous avons présentés se rapportent soit à une diminution du débit de transmission soit à un endormissement. L'endormissement peut être géré par un équipement tiers, un relais de proximité, qui permet la mise en sommeil d'équipements en les déchargeant de certaines tâches. L'inconvénient de ce type de méthode est qu'elle est généralement très dépendante des applications tournant sur les équipements et nécessite de modifier l'architecture existante, avec des éléments qui consommeront de l'énergie. C'est pourquoi nous avons préféré centré nos travaux sur une gestion de l'endormissement par les équipements eux-mêmes.

Pour les deux méthodes, adaptation de débit et mise en sommeil, il est nécessaire de mettre en place une coordination sous forme d'échange d'informations entre les entités appariées du lien. Cet échange s'effectue en bande via le protocole de transfert du lien, ou alors hors bande à partir d'informations d'administrations. Les travaux de la littérature sur le réseau optique et sur Ethernet, indiquent l'intérêt de la méthode d'endormissement devant celle d'adaptation de débit. Cet intérêt réside davantage dans la mise en oeuvre de la méthode que dans le gain obtenu. Ainsi dans l'expérimentation que nous avons réalisée, avec les technologies disponibles à l'instant des travaux, le gain en Ethernet est similaire en utilisant un débit faible à celui obtenu en utilisant un débit important et un mécanisme d'endormissement. Néanmoins l'avantage de l'endormissement réside dans sa capacité à fonctionner quel que soit le support et son débit, le mécanisme est agnostique, ce qui n'est pas le cas pour l'adaptation de débit qui nécessite une signalisation dont la rapidité doit suivre l'augmentation des débits. Nous avons donc orienté nos travaux vers l'utilisation du mécanisme d'endormissement que nous allons évaluer pour un Ethernet standard plus précisément dans le chapitre suivant.

La deuxième partie du chapitre a présenté des travaux avec une stratégie globale qui agissent au niveau du réseau. Il s'agit de mettre en sommeil des liens ou même des routeurs en agissant sur le chemin qui sera suivi par le trafic. Les travaux relèvent du domaine de l'ingénierie de trafic ([Traffic Engineering \(TE\)](#)) qui se rapportent à différentes technologies de réseau : [TE](#) pour [ATM](#), [TE](#) pour [IP](#), [TE](#) pour [MPLS](#) et [TE](#) pour [SDN](#). La technologie [ATM](#) étant en voie de disparition chez de

nombreux opérateurs et n'étant pas utilisée dans les entreprises, nous avons choisi d'étudier les 3 dernières approches. L'approche **IP** fera l'objet du chapitre 3 alors que l'approche **SDN** couplée à une approche **MPLS** sera présentée dans le chapitre 4.

Les mécanismes d'ingénierie visant à économiser l'énergie, s'appuient sur la mise en sommeil d'équipements et peuvent donc poser des problèmes de qualité de service. L'ingénierie peut augmenter les délais de transfert en raison de temps de sommeil et de temps de transferts accrus par des chemins plus longs empruntés par les trafics. Par ailleurs, l'ingénierie verte peut également être source de congestion en concentrant le trafic sur certains liens et donc occasionner des pertes dans le réseau. Dans la suite des travaux nous nous attacherons à examiner l'impact de nos propositions sur la qualité de service.

CHAPITRE 2

Ethernet Efficace en Énergie

Sommaire

2.1	Le standard Energy Efficient Ethernet	40
2.1.1	Intégration du mode de mise en sommeil dans l'architecture en couche	41
2.1.2	Fonctionnement	45
2.2	Implantation EEE et évaluation de consommation	47
2.2.1	Banc de test et caractéristiques de simulation	47
2.2.2	Modèle de puissance	50
2.2.3	Résultats	51
2.3	Amélioration du standard	52
2.3.1	Algorithmes de coalescence	53
2.3.2	Intérêt et problème de la coalescence de trames	56
2.4	Conclusion	58

Nous analysons l'économie d'énergie dans le réseau Ethernet, la technologie cible de nos travaux de thèse. Nous présentons le standard d'économie d'énergie Ethernet, Ethernet Energy Efficient, qui utilise un mode de basse consommation, Low Power Idle, ainsi qu'un moyen d'améliorer l'économie d'énergie du standard grâce au mécanisme de regroupement de paquets, la coalescence. Nous évaluons les gains d'énergie par simulation et expérimentation.

Le chapitre précédent était dévolu aux mécanismes d'économie d'énergie dans les différents segments d'une architecture réseau, dans ce chapitre nous nous concentrons sur l'économie d'énergie dans le réseau Ethernet standard. L'organisme IEEE en charge de la standardisation a spécifié un mécanisme de sommeil que nous allons analyser et évaluer et nous montrerons l'intérêt d'étendre ce mécanisme à un niveau global.

2.1 Le standard Energy Efficient Ethernet

L'IEEE en charge de la standardisation du réseau local 802.3, auquel se conforme le produit Ethernet, a établi de nombreuses versions de documents reflétant les évolutions technologiques (supports, débits ..) et fonctionnelles ([Virtual LAN \(VLAN\)](#), [Class of Service \(CoS\)](#)..) des 30 dernières années. Ethernet dans sa version originale s'appuyait sur un mécanisme d'écoute du support pour détecter la présence ou la non présence d'une émission par une station concurrente. Ce protocole nécessitait de faire la distinction au niveau physique entre 1) un support non alimenté indiquant un problème de raccordement, 2) un support alimenté dit *libre* permettant à une station d'émettre et 3) un support alimenté dit *occupé* permettant de bloquer l'émission d'une station jusqu'à la libération dudit support. C'est pourquoi, en fonctionnement non défectueux, c'est à dire dans les cas 2 et 3, le réseau Ethernet consomme de l'énergie indépendamment des données à transmettre. Ainsi, la version Ethernet 100 Mbps a spécifié l'état de support vide avec la transmission de symbole nommé Idle. Les versions suivantes ont poursuivi dans la même lignée : même si il n'y a pas de données à transmettre, il faut occuper le support en transmettant et donc en consommant de l'énergie.

Lorsque l'on considère le nombre important d'entreprises qui utilisent un système Ethernet, ainsi que le nombre croissant de systèmes Ethernet déployés chez les opérateurs au niveau du réseau d'accès, réseau d'agrégation et même réseau transport, le gaspillage d'énergie s'avérait prégnant. En 2010, le groupe IEEE 802.3az a pris en compte les besoins de diminuer la consommation d'énergie en proposant une version Ethernet Economie en Energie, avec un objectif de consommation de 10% de la consommation nominale. Cette évolution a été intégrée au standard de 2012 et détaillée en [section 6 clause 78] pour des versions Ethernet 100Mbps et au-delà. Les équipements ayant le mécanisme d'économie d'énergie l'indiquent durant la phase d'autonégociation standard de Ethernet. Lorsque

les deux partenaires du lien possèdent ce mécanisme, il est utilisé.

Après études de plusieurs propositions, le mécanisme choisi par la standardisation a été celui de la mise en sommeil devant celui de l'adaptation de débit car la signalisation induite par l'adaptation de débit (à travers des Pulses d'autonégociation) s'est avérée plus coûteuse que celle du mécanisme de sommeil [CRN⁺10].

2.1.1 Intégration du mode de mise en sommeil dans l'architecture en couche

Le standard repose sur le mécanisme d'endormissement à basse consommation LPI (**Low Power Idle (LPI)**) dont le principe est équivalent aux méthodes d'endormissement présentées dans le chapitre précédent et que nous détaillons plus précisément. Nous présentons en premier lieu comment se situe le mécanisme dans l'architecture protocolaire Ethernet et comment il est mis en oeuvre dans la couche physique

1) Mode LPI Couches Physiques et MAC

Le positionnement du mécanisme LPI dans l'architecture Ethernet est indiqué en figure 2.1. Il est géré par une entité nommée client LPI, une signalisation permet au client LPI de signaler à la couche physique et à son partenaire du lien, qu'il va y avoir une interruption de la transmission de données et que les composants peuvent rentrer en mode basse consommation. De même, le client LPI est à même de comprendre par la signalisation que son partenaire a envoyé, une indication de ce type.

Plus précisément, le mode LPI interagit avec la couche physique via les deux primitives de service LP-Iddle request et LP-Iddle indication.

- La primitive LP_Iddle_request demandant un passage en mode LPI (via un paramètre LP_Iddle_request positionné en valeur *assert*) a pour effet de positionner le paramètre d'écoute de porteuse en état ON dans la primitive PLS_CARRIER.indication, ce qui bloque la transmission de trames par la couche MAC. Quand le mode LPI n'est plus demandé (la primitive LP_Iddle_request a le paramètre positionné à *de-assert*), l'encodage intertrame IDDLLE est envoyé puis, après un certain temps le paramètre d'écoute de porteuse est remis à OFF pour

permettre au protocole **MAC** de transmettre ses trames. Le temps est défini en fonction du débit de liaison (cf 1.1.2) pour permettre au partenaire de se préparer à un fonctionnement normal.

- La primitive **LP_Iddle_indication** indique au client **LPI** que le partenaire de la liaison a démarré ou arrêté de signaler le mode **LPI**. Cette information provient de la couche physique et des signaux reçus : un signal de réveil (nommé **wake**) et un signal indiquant un état endormi (**refresh**).

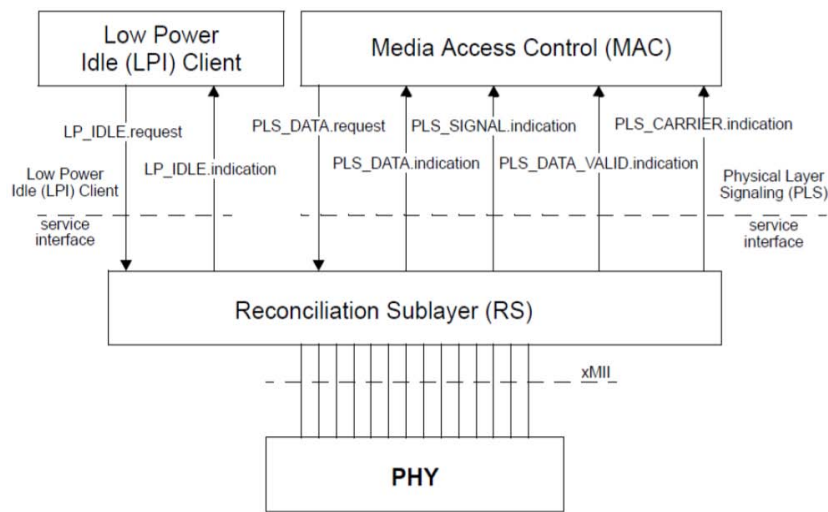


FIGURE 2.1 – Interface entre le client LPI et la couche de réconciliation de Ethernet [IEE12]

Le standard ne se préoccupe pas de savoir quelles sont les conditions pour émettre une primitive **LPI request**, et quelles sont les actions effectuées par le client **LPI**, il indique seulement comment le mécanisme d'économie d'énergie peut être mis en oeuvre.

2) Le mode **LPI** en couche physique

Le pilotage de la mise en sommeil au niveau de la couche physique est illustré en figure 2.2. Le diagramme correspond à un fonctionnement Ethernet à 100 Mbps, il est similaire pour des débits supérieurs.

Le processus de transmission peut désactiver toute ou partie des blocs fonctionnels des sous

couches qui constituent la couche physique. Ils sont nommés **Physical Coding Sublayer (PCS)** **Physical Medium Attachment Sublayer (PMA)** et **Physical Medium Dependant (PMD)**. Lorsque toutes les couches sont désactivées, l'équipement n'est pas en état de recevoir des trames ou encore d'écouter le support et de détecter la présence d'une porteuse.

Le passage en mode **LPI** s'effectue via le positionnement de valeurs :

- sur les signaux d'interface MII à la jonction entre la couche physique et le protocole d'accès (Medium Access Control). Par exemple pour signaler un passage en mode **LPI**, la couche physique va positionner les signaux comme suit : RX_ER en *assert*, RX DV *de-assert*, et RXD <3 :0>.
- sur les signaux de transmission reçus/émis par la sous couche **PMD** (*ex tx quiet, tx-code*).

Plus précisément, les signaux *quiet*, *refresh*, *wake* sont transmis entre émetteur et récepteur. Le standard prévoit un réveil cyclique du processus de transmission pour transmettre un signal de rafraîchissement (*refresh* , à travers les bits *tx_code*) de façon à permettre au récepteur distant de gérer les dérives d'horloges entre les équipements partenaires du lien. Le cycle de rafraîchissement continue jusqu'à la réception de l'encodage normal inter trames par l'interface MII. La fonction physique prévient alors son partenaire qu'il doit se réveiller par un signal *wake* transmis durant un temps prédéfini (dont les valeurs sont spécifiées par technologie de médium).

Nous synthétisons le fonctionnement du mode basse consommation (Low Power Iddle) en figure 2.3.

Le mécanisme d'endormissement est défini à partir de 3 états : réveil (Wake), actif (actif) endormissement (sleep). Les alternances entre mode actif et mode **LPI** à la fois émetteur et récepteur ne sont pas instantanées. Le premier et dernier état sont des états transitionnels qui reflètent les temps pris par les équipements pour s'endormir et se réveiller. Leurs durée est définie dans le standard respectivement par les temps T_w (temps nécessaire pour revenir au mode actif) et T_s (temps nécessaire pour entrer en mode **LPI**). La politique de réveil est définie par l'arrivée d'un paquet à transmettre. Lorsqu'il n'y a pas de données à transmettre, l'émetteur s'endort, il passe en mode **LPI**.

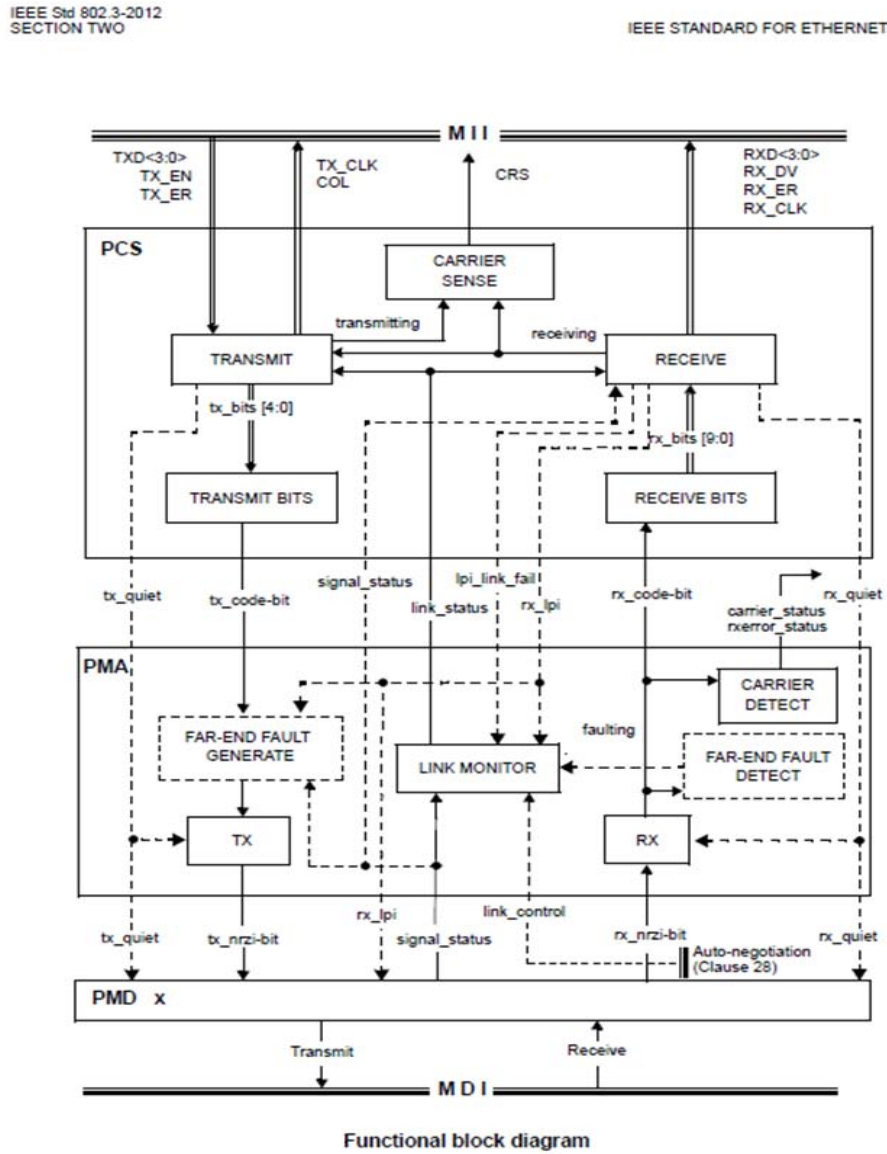


FIGURE 2.2 – Architecture standard des blocs fonctionnels de la couche physique Ethernet [IEE12]

Ce mode n'est pas borné dans le temps mais l'émetteur endormi se réveille de façon cyclique, toutes les T_q secondes, pour transmettre une information durant un temps T_r , nommée Refresh sur la figure 2.3, elle permet la synchronisation avec le récepteur.

Les valeurs des temps T_w et T_s sont indiquées par des bornes minimales et maximales selon les

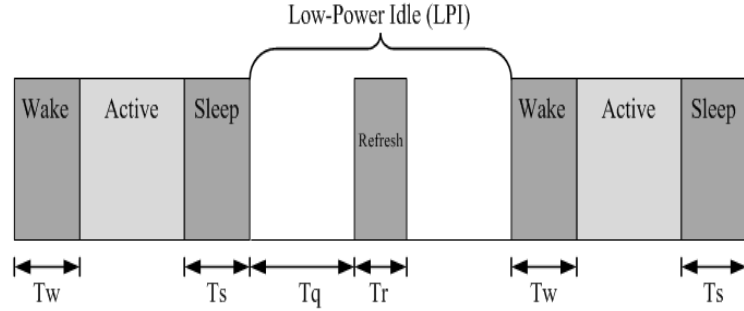


FIGURE 2.3 – Mode basse consommation Ethernet standardisé (LPI : Low Power Idle)

débits associés à la version Ethernet, à savoir 100Mbps, 1Gbps ou 10Gbps (tableau 2.1). Avec pour valeurs min de T_w et T_s pour les liaisons 1000Base-T respectivement 16.5 s et 182 s, [eee10]. Nous utiliserons ces valeurs pour simuler le fonctionnement de [Energy Efficient Ethernet \(EEE\)](#) dans la suite de ce chapitre.

TABLE 2.1 – Paramètres du mode LPI par version de débit Ethernet

Protocol	Tw (s)	Ts (s)	
	<i>Min</i>	<i>Min</i>	<i>Max</i>
<i>100Base-TX</i>	<i>30</i>	<i>200</i>	<i>220</i>
<i>1000Base-T</i>	<i>16.5</i>	<i>182</i>	<i>202</i>
<i>10GBase-T</i>	<i>4.48</i>	<i>2.88</i>	<i>3.2</i>

2.1.2 Fonctionnement

A partir du descriptif du standard que nous venons de présenter, nous élaborons le fonctionnement de l'EEE sous forme de machine à états finis conformément à la Figure 2.4 que nous pouvons implanter dans un simulateur.

Les états sont représentés par des lignes verticales. Le nom de l'Etat est mentionné au-dessus de l'état (à savoir, [LPI](#), Wake, Active Sleep, Refresh)

- Les transitions sont représentées par des flèches de l'état source à l'état destination.

Précisons, les cinq états. Nous utilisons la terminologie anglaise du standard

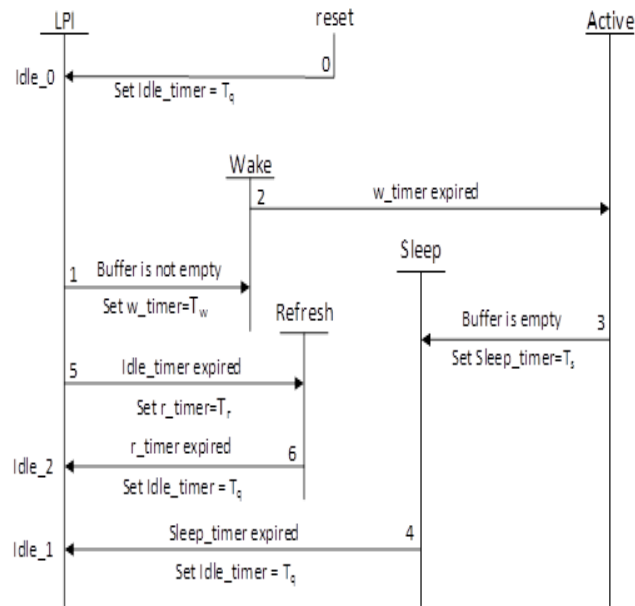


FIGURE 2.4 – Machine à état finis de la fonctionnalité de mise en sommeil Ethernet Standard

- Actif : Le lien est entièrement alimenté et opérationnel. Les paquets sont en attente dans la mémoire tampon et transmis.
- Wake : Le lien est en cours de démarrage.
- Sleep : Le lien est en sommeil.
- Refresh : Le lien est en cours d’actualisation.
- Idle : Le lien est en faible puissance ou en mode veille

Quatre temporisateurs sont définis pour gérer les temps passés dans les 4 derniers états : `w_timer`, `Sleep_timer`, `r_timer`, `idle_timer`. Ils sont respectivement initialisés avec les valeurs T_w , T_s , T_r et T_q (Time quiet). Lorsque la valeur du timer atteint zéro, une transition s’ensuit.

- Transition 0 : Un lien démarre en mode **LPI**. Le timer `idle` est armé à la valeur T_q
- Transition 1 : Lorsque les paquets arrivent au tampon, le tampon n’est pas vide. La liaison se réveille pour transmettre le paquet. Le timer T_w est armé à T_w .
- Transition 2 : Expiration du timer T_w . Le lien entre en état actif et les paquets sont envoyés.

- Transition 3 : Lorsque la liaison a transmis tous les paquets dans le tampon, le tampon est vide. Le timer T_{sleep} est armé à T_s , la liaison se prépare à rentrer en sommeil.
- Transition 4 : Expiration du sleep timer. Armement du Idle timer, passage à l'état iddle.
- Transition 5 : Expiration de I_{idle_Timer} . Passage à l'état refresh, armement du r_timer .
- Transition 6 : Expiration du r_timer . Passage à l'état Iddle.

L'implantation de cet automate dans le simulateur NS3 ([Network Simulator 3 \(NS3\)](#)), va nous permettre d'évaluer l'intérêt du standard.

2.2 Implantation EEE et évaluation de consommation

Nous avons présenté le fonctionnement du standard et les moyens de le modéliser par simulations. Nous allons à présent évaluer ses performances par simulation et les confronter à des mesures réelles obtenues par expérimentation. Nous présentons le banc de test et analysons les résultats de mesure énergétique que nous avons obtenus.

2.2.1 Banc de test et caractéristiques de simulation

Nous avons élaboré le banc de test représenté en Figure 2.5 ;

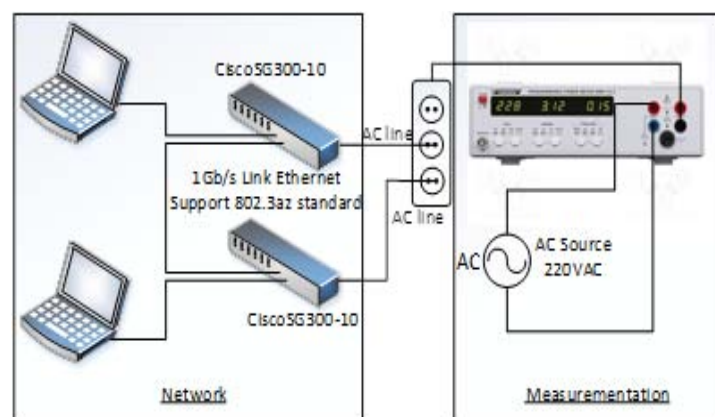


FIGURE 2.5 – Banc d'essai IEEE802.3az

Il est constitué de :

- **Commutateurs** : deux SG-300-10 commutateur Cisco [cis] (version matérielle 1, version de logiciel 1.2.7.76). La norme IEEE802.3az est présente sur toutes les interfaces (SG300-10 a au maximum dix ports actifs).
- **Liens Ethernet** : Chaque ordinateur personnel (PC) est connecté au Cisco SG300-10 par un lien Gigabit Ethernet (1000Base-T). Les deux commutateurs Cisco sont également reliés entre eux par un lien Gigabit Ethernet. Tous les câbles ont des longueurs inférieures à 5 mètres.
- **Flux de trafic** : Deux PC de test, sont utilisés en tant que source et destination de trafic IPv4. Le trafic est crée par le produit [Distributed Internet Traffic Generator \(D-ITG\)](#) [EPV04]. D-ITG est une plate-forme capable de générer du trafic IPv4 [AGE⁺04] [EPV04] [BDP10] [BDP12]. La Figure 2.6 illustre l'architecture D-ITG. ITGSend est le composant chargé de générer du trafic vers le composant ITGRecv. En exploitant une conception multithread, ITGSend peut envoyer en parallèles plusieurs flux vers plusieurs instances ITGRecv et le composant ITGRecv peut recevoir du trafic découlant de plusieurs instances de ITGSend. Un canal de signalisation est créé entre chaque paire de composants ITGSend et ITGRecv pour contrôler la génération des flux. ITGSend et ITGRecv peuvent éventuellement produire des fichiers journaux contenant des informations détaillées sur tous les paquets envoyés et reçus qui seront analysés par le composant ITGDec.
- **Ordinateur personnel (PC)** : Les ordinateurs ont été équipés chacun d'une carte qui prend en charge la norme 802.3az. L'ordinateur source est connecté au commutateur de source et l'ordinateur de destination connecté au commutateur de destination. De façon à intégrer facilement le générateur de flux nous avons adopté le système d'exploitation Linux pour notre expérience.

Les conditions de simulations sont données dans le tableau 2.1, 2.2, 2.3. Y sont indiqués, les paramètres des algorithmes simulés et le modèle de puissance.

- Les paramètres de trafic sont similaires à ceux utilisés en [GS03] et [HARPFVLG11] qui utilisent des tailles constante de paquets de 1000bits avec une arrivée de paquets selon un processus de poisson. Nous faisons varier la charge de la liaison de 0 à 100%.
- Les valeurs des timers sont conformes au standard. Les valeurs de puissance (Tables 2.2 et 2.3)

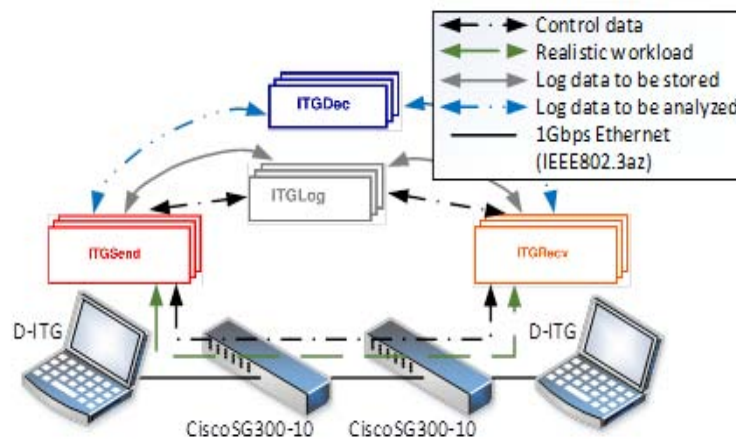


FIGURE 2.6 – Architecture de mesure

proviennent de la littérature, elles ne sont pas réalistes mais cohérentes. Nous aborderons un aspect plus réaliste dans le paragraphe dédié aux résultats d'expérimentation en [GS07a].

- Les valeurs des timers sont conformes au standard (Table 2.1)
- En ce qui concerne les valeurs de puissance, mentionnons que les valeurs de puissance proviennent de la littérature, et que bien que non réalistes comme nous l'avons expérimenté, elles sont cohérentes. Comme nous le montrerons dans le paragraphe dédié aux résultats d'expérimentation (L'analyse de l'importance des valeurs des puissances de l'algorithme que nous avons publiée en [PDJC13] indique que les gains d'énergie sont fonction de ces valeurs tout en gardant un comportement de variation similaire).

Plus précisément nous considérons, pour l'émetteur et le récepteur, les puissances consommées par :

- les transitions entre états actifs et endormis (Tx_Down, Rx_Down) et inversement entre état endormi et état actif (Tx_Up, Rx_Up).
- la transmission, réception de paquet en état éveillé (Tx_Busy, Rx_Busy).
- la non transmission/réception d'information en état éveillé (Tx_Ready, Rx_Ready). L'équipement ne traite pas de trame mais son temps d'activité n'est pas écoulé (algorithme de coalescence).

- le sommeil (Tx_Sleep, Rx_Sleep).

TABLE 2.2 – Paramètres de consommation en transmission

Transmitter	Power Level (Watt)
Tx_Ready	1
Tx_Busy	2
Tx_Sleep	0.1
Tx_Up	2
Tx_Down	2

TABLE 2.3 – Paramètres de consommation en émission

Receiver	Power Level (Watt)
Rx_Ready	1
Rx_Busy	1.3
Rx_Sleep	0.1
Rx_Up	1.3
Rx_Down	1.3

2.2.2 Modèle de puissance

Nous définissons le modèle de puissance du banc de test pour l'équipement S de N ports, comme suit.

$$P(S) = P(chassis) + N e \quad P.port_endormi + \sum_{i=1, Na} P.port_actif(i); N_e + N_a = N \quad (2.1)$$

La puissance consommée par un équipement switch, P(S) est constituée de 3 composantes :

1. P(chassis) : c'est la puissance mesurée sur un équipement alimenté qui n'effectue aucune action.
2. P.port_endormi : c'est la puissance consommée par un port connecté mais qui ne supporte pas de trafic, il est dans l'état endormi si il supporte le standard et que l'équipement appairé le supporte aussi, sinon il est en état Iddle. Il est supposé une configuration de Ne ports endormis

3. $P_{\text{port_actif}}(i)$: c'est la puissance consommée par le port i supportant une charge de trafic τ_i

Le nombre de ports actifs est de N_a avec $N_a + N_e = N$ le nombre total de ports de l'équipement.

La détermination des éléments du modèle se fait de la manière suivante. La puissance du chassis s'obtient en mesurant par analyseur la puissance de l'équipement avec tous les ports non connectés. Nous avons expérimenté sur différents modèles de constructeurs et obtenus des valeurs différentes. Pour s'abstraire de ce problème, nous allons par la suite considérer l'intérêt d'un mécanisme d'énergie en mesurant le gain obtenu.

La puissance associée à un port en état endormi est obtenue en mesurant la puissance des 2 switchs raccordés entre eux mais avec tous les autres ports non raccordés, et en divisant par 2 cette valeur.

Pour obtenir la valeur de puissance d'un lien actif il faut mesurer le banc d'essai en générant du trafic et en retranchant les valeurs déterminées précédemment.

2.2.3 Résultats

Sur les résultats de la figure 2.7, nous constatons un fonctionnement de l'économie d'énergie similaire entre celui défini dans le standard que nous avons modélisé et introduit dans le simulateur NS3 avec celui implanté réellement chez le constructeur.

Nous notons cependant des différences qui s'expliquent par des valeurs de consommation différentes. Celles de la simulation correspondent à des valeurs prises dans la littérature et non des valeurs réelles. Ces valeurs réelles que nous présentons ici sont celles de CISCO elles sont différentes de celles de DLink [Dli], et correspondent à un modèle précis d'équipement. Il est alors difficile de les utiliser pour développer un modèle de coût énergétique qui pourrait être employé par un administrateur de réseau souhaitant mettre en place une politique d'ingénierie verte.

Un moyen de résoudre ce problème serait de rajouter au niveau des protocoles d'administration spécifique à Ethernet ([Link Layer Discovery Protocol \(LLDP\)](#)) ou plus généraux (comme [Simple Network Management Protocol \(SNMP\)](#)) des informations sur le modèle de consommation de chaque équipement. Par ailleurs, une gestion centralisée du plan de contrôle permettrait de définir pour chaque constructeur les coûts de consommation des équipements.

Conclusion : la simulation et l'expérimentation indiquent une courbe de coût de consommation énergétique semblable, fonction de la charge des liens avec des divergences sur les valeurs. Dans cette thèse nous utiliserons les valeurs de la littérature et plutôt que de mesurer la consommation d'énergie nous évaluerons le gain d'énergie. Nous testerons sur banc d'expérimentation nos propositions afin de vérifier la cohérence des résultats. Nous retiendrons de ces résultats l'importance de la charge dans la consommation et la croissance très rapide de la consommation en fonction de celle-ci

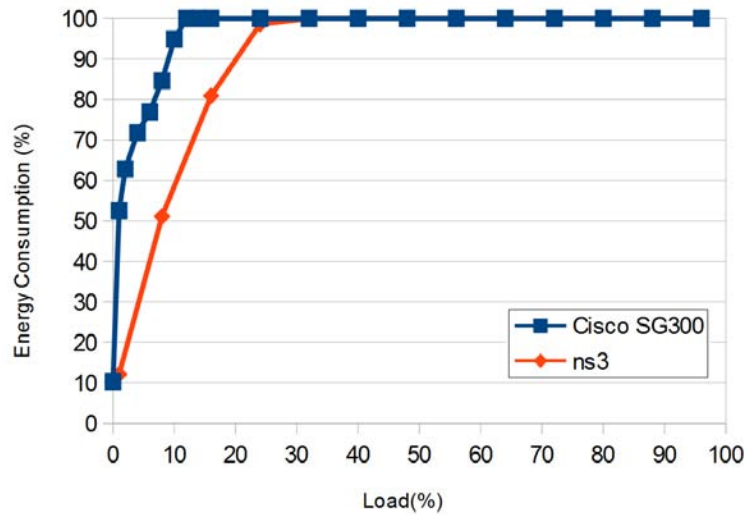


FIGURE 2.7 – Comparaison de la consommation d'énergie IEEE802.3az entre banc d'essai et simulation

2.3 Amélioration du standard

[MC12] [CRN⁺10] proposent d'améliorer le standard Ethernet par un mécanisme de coalescence qui augmente le temps de mise en sommeil et diminue la consommation d'énergie du standard. Nous allons détailler ce mécanisme avant d'en évaluer par simulation l'intérêt en terme d'économie d'énergie réalisée en regard de l'utilisation du standard [EEE](#).

2.3.1 Algorithmes de coalescence

Le principe de la coalescence intervient avant l'émission de trame. Il s'agit de mémoriser dans un tampon, les trames devant être émises alors que le transmetteur est en état de veille, pour lui permettre de rester dans cet état.

Plusieurs algorithmes de réveil ont été proposés. Gupta Singh [GS03] dans leur mécanisme de mise en sommeil, étudient la mise en file d'attente de paquets avant émission. Ils supposent que l'arrivée des paquets dans une interface suit localement un processus de poisson sur une courte période. Ils en estiment la moyenne à court terme en mesurant le temps séparant l'arrivée de deux paquets (en pratique les 5 derniers paquets). Lorsque l'occupation de la file, q , est inférieure à un seuil, b , ils estiment la moyenne d'arrivée pour prédire le temps, t_s , pendant lequel l'interface peut se mettre en sommeil sans que la file d'attente ne dépasse une valeur maximale de b , et communiquent cette valeur d'endormissement au récepteur (leur mécanisme qui a été publié avant la standardisation n'utilise pas la signalisation du standard). Au bout de ce temps l'interface se réveille et transmet les paquets. Si la file se retrouve vide, un intervalle de sommeil de même valeur que le précédent démarre. De façon à avoir une probabilité faible de dépasser le seuil, ils suggèrent de fixer la valeur de b à 10% de la capacité totale de la file (la formulation analytique du problème sera trouvée en [GS07a] [HARPFVLG11]). Nous retenons leur préconisation pour la valeur de b .

Pour éviter d'avoir un délai trop important sur les paquets et pour éviter d'avoir recours à un calcul de moyenne qui ne s'exprime pas sous forme produit mais nécessite des approximations [HARPFVLG11] propose d'utiliser l'algorithme *zero backlog*. Son objectif est de dormir moins souvent mais plus longtemps que dans l'algorithme précédent. Ce qui augmente l'économie d'énergie par rapport à l'algorithme de Gupta puisque l'on diminue l'énergie consommée pendant les phases transitoires (énergie non prise en compte par Gupta et al.). L'algorithme *zero backlog* ne déclenche la période de sommeil que lorsque la file d'attente est complètement vide. Cette période de sommeil est limitée par un temps de timer pour éviter des délais trop importants. Ce temps est calculé pour minimiser l'énergie, il doit être suffisamment grand pour que l'énergie consommée par le réveil et l'endormissement soit inférieure à celle consommée par une interface réveillée qui ne transmet pas.

L'algorithme *zero backlog* met en sommeil la carte lorsque le tampon est vide c'est-à-dire que

le trafic est faible et que donc un endormissement devrait se produire. Le deuxième algorithme également proposé par [HARPFVLG11] est *le lazy start algorithm*, il a pour but de créer des temps de sommeil. Le timer qui limite l'endormissement ne commence à décompter le temps que lorsqu'il y a un paquet à transmettre. L'algorithme propose également de se réveiller sur un seuil d'occupation pour éviter de se réveiller à chaque fois qu'il y a quelque chose à transmettre au risque de perdre de l'énergie avec les phases de transition. Nous avons choisi d'évaluer plus précisément cet algorithme car il permet de limiter le délai par rapport à celui de Gupta et il offre de meilleures performances que le zero backlog [HARPFVLG11].

L'algorithme *Lazy start* définit une mémorisation qui s'arrête sur un seuil de remplissage du tampon (variable COUNT sur la figure) ou sur un temps (TIMER). L'avantage de cette méthode est comme l'illustre la figure 2.8 de créer des intervalles de temps permettant à la liaison de passer en mode LPI. Dans cette figure le débit de transmission de l'émetteur est supposé constant avec un temps inter trames trop court (inférieur à T_s) pour permettre un passage en LPI. Le nombre minimum de paquets à partir duquel l'interface s'active, qw , vérifie l'équation suivante :

$$p_a \cdot count \cdot C = p_a \cdot T \quad (2.2)$$

D'où $count = \frac{C \cdot T}{p_a}$.

Où C est le débit de l'interface, p_a : la puissance consommée en état actif et T le temps de transition entre l'état sommeil et l'état actif.

Grâce au regroupement de trames, il est possible de passer en état sommeil alors que ce n'eût pas été possible sans ce regroupement. Cependant cette méthode présente un inconvénient inhérent à la mémorisation, il y a une augmentation du délai qui pourrait être préjudiciable à des trafics à délais contraints. Nous évaluons ce point dans le chapitre suivant.

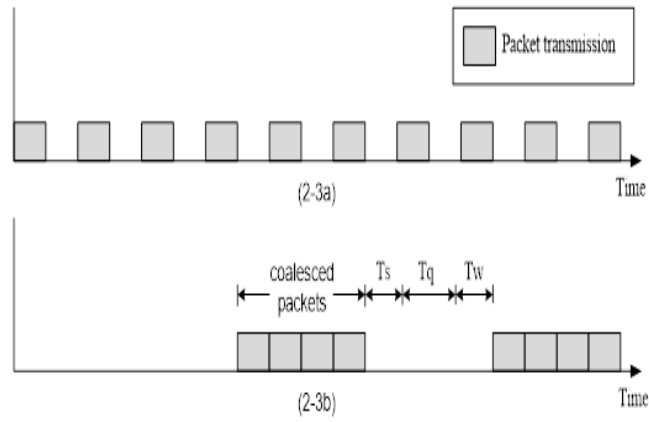


FIGURE 2.8 – Mécanisme de coalescence : Transmission de paquets sans coalescence (2-3a) avec coalescence (2-3b) [MC12]

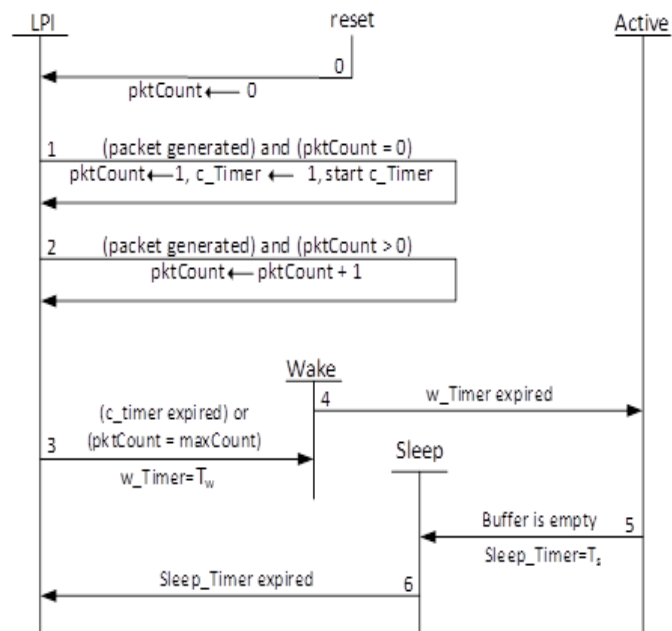


FIGURE 2.9 – Fonctionnement du standard EEE avec coalescence

2.3.2 Intérêt et problème de la coalescence de trames

Pour évaluer les avantages et inconvénients du mécanisme de coalescence, en regard du standard [EEE](#) nous avons implanté l'algorithme décrit en figure [2.4](#) dans le simulateur [NS3](#) [[NS3](#)]. Les états et transitions sont ceux présentés précédemment en Figure [2.9](#) complétés par les variables de l'algorithme de coalescence indiqués en Table [2.4](#). Sont rajoutés aux timers du standard (sleep_Timer et w_timer), le timer qui contrôle le temps de mémorisation des paquets (c_timer), ainsi qu'un compteur (pk_count) pour contrôler que le nombre de paquets mémorisés est inférieur à un seuil nommé Max_count. Nous avons conservé la limite de 10% préconisée par [[GS03](#)] pour le seuil et assigné une valeur de tampon à 300. Cette valeur a été fixée à partir de résultats de simulation montrant que l'augmentation de la taille du tampon au-delà de 300 n'améliore pas les performances énergétiques. En effet le temporisateur armé à l'arrivée du premier paquet expire presque toujours avant que la condition de seuil ne soit réalisée. De même nous avons testé différentes valeurs de timers et de paquets. La valeur de 2.75ms utilisée en [[HARPFVLG11](#)] correspond à une proportion similaire d'événement de mise en sommeil par timer et de mise en sommeil sur occupation de buffer. Les résultats et courbes de notre analyse ont été publiés en [[PDJC13](#)].

Le scénario d'évaluation est implanté sur le simulateur ([NS3](#)) entre deux noeuds, un émetteur et un récepteur reliés par un lien Ethernet 1000base-T. Deux fonctionnements sont simulés.

Cas 1 : L'émetteur et le récepteur utilisent un Ethernet efficace en énergie avec l'algorithme de la figure [2.4](#). Dans cet état la carte réseau est complètement en veille sans possibilité d'écouter le lien.

Cas 2 : l'émetteur utilise un algorithme [EEE](#) coalescent. Les paquets sont stockés en file d'attente pour un certain temps. Les paramètres clés de l'algorithme sont Qmax, et Tmax avec Qmax la taille du tampon de stockage correspondant à la capacité de file d'attente utilisée dans le simulateur et Tmax la valeur du timer de coalescence.

La figure [2.10](#) illustre l'intérêt de la coalescence. Ce mécanisme permet d'obtenir un gain d'énergie pour une charge de la liaison pouvant aller jusqu'à environ 50% alors qu'avec le standard le gain d'énergie est limité à des charges inférieures à 25%. Il augmente également l'économie d'énergie réalisée à faible charge (22% de gain de plus qu'avec le standard, pour une charge de liaison de 25%)

TABLE 2.4 – Paramètres de simulations

	EEE	EEE with packet coalescing
Traffic type	Poisson	Poisson
Packet Size	1000 bytes	1000 bytes
Number of Packet	100000 packets	100000 packets
Qmax	non	300 packets
Tmax	non	0.00275 second
Count-max	non	10% of Qmax

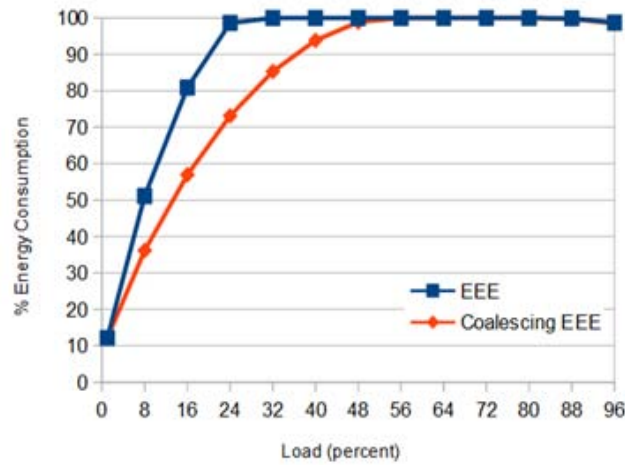


FIGURE 2.10 – Consommation d'énergie de l'IEEE 802.3az avec et sans coalescence

Cependant le mécanisme de coalescence n'est pas standardisé, il n'est pas forcément disponible sur les équipements. En outre, il est difficile en pratique de le configurer (valeur de T_{max} , Q , $b...$). C'est pourquoi nous proposons d'améliorer le standard en utilisant une approche d'ingénierie de trafic effectuée au niveau réseau.

2.4 Conclusion

Dans ce chapitre nous avons analysé et évalué l'intérêt du standard d'économie d'énergie par un mécanisme d'endormissement de liens en technologie Ethernet, puis nous nous sommes intéressés au moyen de l'améliorer.

Nous avons présenté le fonctionnement du standard sous forme d'automates que nous avons implantés dans un simulateur et avons comparé les résultats de simulation avec des résultats d'expérimentation. Les résultats que nous avons obtenus mettent en avant la cohérence de la simulation et de l'expérimentation tout en faisant ressortir la difficulté d'obtenir un modèle de puissance générique pour le standard qui ne soit pas dépendant d'un constructeur. Dans la suite de cette thèse nous utiliserons des valeurs issues de la littérature, qui nous permettront d'obtenir des comparaisons de solutions pertinentes par simulation.

L'évaluation du standard nous a permis de mettre en avant l'importance de la charge sur la consommation et nous nous proposons d'établir un modèle de puissance intégrant la valeur de la charge.

Le standard Ethernet permet des gains d'énergie à faible charge (inférieure à 25%) qui peuvent être augmentés en créant des périodes de sommeil. Dans ce chapitre nous avons indiqué des algorithmes de coalescence et évalué sur l'algorithme lazy start un gain pouvant atteindre jusqu'à 35 % pour une charge donnée. Cependant il n'existe pas d'algorithme standardisé et le réglage des paramètres de l'algorithme est complexe. C'est pourquoi nous proposons d'utiliser un mécanisme global au niveau du réseau pour créer des périodes de sommeil, et développons dans le chapitre suivant notre proposition d'ingénierie verte.

En s'appuyant sur les résultats d'évaluation que nous avons obtenus, nous allons chercher à améliorer l'économie d'énergie réalisée par le standard en prenant en compte la charge des liens pour diriger le trafic.

CHAPITRE 3

Ingénierie verte distribuée

Sommaire

3.1	Ingénierie de trafic verte adaptée à la technologie EEE	60
3.1.1	Principaux éléments OSPF	61
3.1.2	Modèle de coût OSPF pour réseau Ethernet Efficace en Energie . .	62
3.1.3	Illustration des politiques d'ingénierie verte proposées	65
3.2	Evaluation du fonctionnement de l'ingénierie par simulation et expérimentation	68
3.2.1	Evaluation par simulation	69
3.2.2	Evaluation expérimentale	77
3.3	Comparaison de la proposition sur un scénario réaliste	81
3.3.1	Comparaison à une solution de mesure idéale	81
3.3.2	Comparaison à la solution optimale Green-TE	86
3.4	Architecture multicouche - Prise en compte de l'énergie lien	90
3.4.1	Proposition de métrique verte en fibre optique	90
3.4.2	Evaluation	92
3.5	Conclusion	97

Notre objectif est d'étendre les capacités d'économie d'énergie du standard [EEE](#) étudié dans le chapitre précédent qui est un mécanisme de niveau lien local à un niveau global, le réseau internet. Nous proposons une ingénierie de trafic [IP](#) verte qui repose sur un routage économe en énergie. L'infrastructure réseau que nous étudions suppose une couche de transfert qui utilise la norme IEEE 802.3 [EEE](#) et une pile protocolaire [IP](#). Nous l'étendons à une architecture multicouches optique/Ethernet/[IP](#).

Comme nous l'avons vu dans le chapitre précédent le mécanisme d'endormissement permet de verdir le lien Ethernet, dans ce chapitre nous proposons d'étendre la capacité de verdissement par un plan de contrôle réseau adapté. Nous nous inscrivons dans des mécanismes standards de telle sorte que l'approche soit aisément déployable. Nous allons d'abord décrire notre proposition puis nous en évaluerons l'intérêt au moyen de simulation et expérimentation. Nous terminons ce chapitre en étendant notre approche à un réseau multicouches, optique et internet.

3.1 Ingénierie de trafic verte adaptée à la technologie EEE

L'ingénierie de trafic permet de contrôler le chemin suivi par les paquets IP dans le réseau. Dans un objectif de verdissement elle répartit le trafic sur les chemins du réseau en sélectionnant des routes qui ne l'eussent pas été par un calcul de plus court chemin conventionnel reposant sur un calcul de plus court chemin en nombre de noeuds, à la capacité des liens. Dans ce chapitre nous utilisons l'approche d'ingénierie IP. Nous étudierons les approches MPLS/SDN dans le chapitre suivant.

La politique d'ingénierie que nous proposons privilégie le passage par des liens qui ne se mettent pas en basse consommation. Pour ce faire, nous définissons un calcul des chemins effectué à partir d'un modèle de coût des liens reflétant le passage en mode basse consommation. Par rapport aux travaux existants dans le domaine nous proposons d'utiliser un modèle de coût adapté à la technologie Ethernet.

La question que nous traitons est comment définir le modèle de coût vert pour répartir le trafic de façon à permettre aux liens Ethernet de dormir le plus possible. Nous ne cherchons pas à ne pas transmettre du tout sur les liens, ce qui correspondrait à un modèle de coût de type ON/OFF où un lien inutilisé est privilégié au détriment d'un lien utilisé même faiblement, mais nous cherchons à transmettre une quantité de trafic suffisamment faible pour que le lien Ethernet puisse passer en mode sommeil.

Nous proposons d'utiliser de nouvelles métriques de routage prenant en compte la charge des liens en se référant au modèle de puissance de la technologie Ethernet étudiée dans le chapitre 2, que nous intégrons dans le protocole de routage standard Internet Engineering Task Force (IETF) de l'internet : OSPF [rfc98]. C'est une approche distribuée où chaque routeur se réfère à sa table

de routage pour diriger le trafic. Par ailleurs nous proposons plusieurs politiques d'ingénierie pour prendre en compte l'énergie mais également la qualité de service, en évitant la congestion.

Nous présentons à la suite succinctement [OSPF](#) puis les métriques que nous proposons.

3.1.1 Principaux éléments OSPF

Open Shortest Path First ([OSPF](#)) est un protocole de routage à état de liens. Les principales fonctions de [OSPF](#) sur chaque routeur sont 1) l'émission en diffusion à destination des autres routeurs du domaine des états du routeur, c'est-à-dire des caractéristiques de ses interfaces de raccordement (à qui il est connecté et par quelle technologie) 2) la réception des états de liens envoyés, 3) l'interprétation de ces états pour en déduire le graphe de la topologie et le coût associé à chaque arrête 4) le calcul sur le graphe du plus court chemin (algorithme de Dijkstra) et finalement 5) la construction des tables de routage indiquant l'interface de sortie vers laquelle diriger un paquet qui a une adresse destination donnée.

Nous allons modifier le routage [OSPF](#) standard en agissant sur la fonction 1). Nous enrichissons l'état des informations échangées par le routeur, nommées LSA ([Link State Advertisements \(LSA\)](#)).

Link State Advertisements (LSA) : Le [LSA](#) est la mise en oeuvre du mécanisme d'annonce utilisé par [OSPF](#) pour fournir des valeurs d'états de liaison, c'est-à-dire des valeurs de coûts [OSPF](#) des liaisons entre le routeur émetteur et ses voisins. Les valeurs de coût sont échangées de façon cyclique afin de prendre en compte des changements de topologie et pouvoir calculer un nouveau chemin en cas de panne. Les routeurs conservent une copie du dernier échange dans la base de données d'état de liaison ([Link State Database \(LSDB\)](#)).

Le coût est généralement déterminé par une équation telle :

$$Coût_{OSPF} = \frac{10^8}{Bandwidth} \quad (3.1)$$

L'objectif du choix de route est de privilégier la performance du point de vue du délai. Plus la bande la bande passante du lien est élevée plus le coût associé est faible, le plus court chemin étant

calculé à partir de ces coûts il tendra à passer par des liens à coûts faibles et donc rapides.

Shortest Path First : Chaque routeur dans le système autonome est représenté comme une racine d'un arbre où les feuilles sont les destinations. Le chemin de la racine à l'arbre est un plus court chemin calculé avec les coûts [OSPF](#) de la base de données d'état de liaison ([LSDB](#)).

Table de routage : Les routeurs se réfèrent à la table de routage pour connaître l'interface de sortie sur laquelle émettre un paquet entrant.

3.1.2 Modèle de coût OSPF pour réseau Ethernet **E** cace en Energie

Nous proposons de modifier l'équation de coût [OSPF](#) pour refléter au mieux le fonctionnement d'un lien Ethernet capable de se mettre en sommeil, de façon à éviter de diriger le trafic vers des liens susceptibles de se mettre en sommeil. Les autres éléments de [OSPF](#) restent inchangés.

Comme indiqué au chapitre 2, l'économie d'énergie réalisée avec l'IEEE 802.3az est importante à très faible charge de trafic et diminue avec l'augmentation de la charge de trafic comme le montre la figure [3.1](#). Dans un réseau de trafic élevé, l'émetteur transmet quasi en permanence des données au récepteur, ce qui empêche les deux éléments d'accéder au mode de [LPI](#), il n'y a pas de gain d'énergie par rapport à un transfert Ethernet classique. Au contraire, dans un réseau de trafic faible, l'émetteur qui n'a pas à transmettre durant des périodes suffisamment longues en regard des temps définis dans le standard peut avec le récepteur entrer fréquemment en mode basse consommation et induire une économie d'énergie.

Une façon simple de refléter ce fonctionnement est d'utiliser un modèle de coût à deux états : économie, non économie. Au-delà d'un seuil de charge le coût est élevé, pour refléter la non économie d'énergie induite par une charge importante de trafic, en deçà le coût est faible : comme il peut y avoir économie d'énergie, il y a intérêt à diriger le trafic sur ce lien.

L'équipement Ethernet selon son constructeur peut avoir sa propre valeur de seuil il est alors possible de positionner par le protocole d'administration idoine ([LLDP](#) , [SNMP](#) ou [Command Line](#)

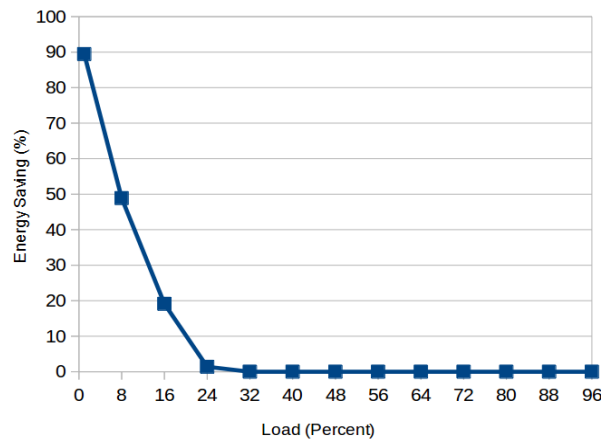


FIGURE 3.1 – Economie d’énergie du standard IEEE 802.3az à différentes charges de trafic

[Interface \(CLI\)](#)) la valeur du seuil sur chaque équipement. L’intérêt de ce modèle est de pouvoir simplement refléter la diversité de consommation des équipements selon leurs constructeurs en modifiant la valeur de seuil par un paramétrage administrateur. Par défaut, la valeur du seuil est celle correspondant à 25% de charge en référence à la figure 3.1.

Nous nommons [EAGER](#) la métrique de routage additive, correspondant à cette modélisation, que nous formalisons par :

“EAGER” or Extremely AuGmented Energy efficient etheRnet

Étant donné le lien (i, j) associé à une valeur d’utilisation de liaison $L_{i,j}$, la valeur [EAGER](#) de ce lien est calculée par :

$$EAGER_{i,j} = \begin{cases} 100 & , \text{if } L_{i,j} \leq 25 \\ 1 & , \text{Otherwise} \end{cases} \quad (3.2)$$

Une valeur de 100 est associée à un lien faiblement chargé de façon à diriger le trafic vers un lien plus chargé. Sinon le coût est de 1, cette faible valeur fera choisir préférentiellement ce chemin dans le calcul de plus court chemin (Figure 3.4).

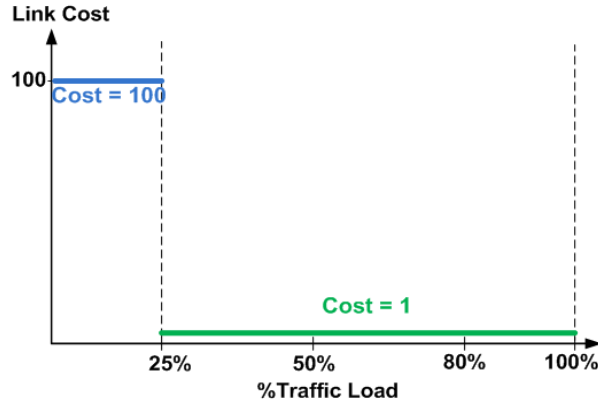


FIGURE 3.2 – Représentation graphique de EAGER

“CARE” or Congestion aware Augmented eneRgy efficient Ethernet

Une modélisation plus élaborée du fonctionnement [EEE](#) consiste à suivre plus finement l'économie d'énergie réalisée selon la valeur de la charge du lien.

Alors que dans la métrique Eager l'économie d'énergie prise en compte n'est pas fonction de la valeur L_{ij} d'utilisation du lien, nous introduisons ce paramètre dans la métrique que nous nommons [CARE](#). Par ailleurs de façon à ne pas trop surcharger des liens au risque de générer des pertes par congestion, nous introduisons une valeur de seuil de congestion : X .

La métrique [CARE](#) est définie comme suit.

$$CARE \quad X_{i,j} = \begin{cases} \frac{100}{L_{i,j}} & , \text{if } L_{i,j} \leq 25 \\ 1 & , \text{if } 25 < L_{i,j} \leq X \\ 100 & , \text{if } L_{i,j} > X \end{cases} \quad (3.3)$$

La valeur du seuil de congestion est apte à refléter différentes exigences de service, telle, la volonté de ne pas surcharger le lien de façon à ne pas avoir un délai trop important ou encore, la volonté de ne pas faire passer tout le trafic sur un lien pour en cas de panne perdre peu d'information.

Nous allons étudier deux valeurs de seuil, à savoir :

$X=0.5$ qui correspond aux valeurs de [Maximum Link Utilization \(MLU\)](#) considérées dans de nombreux algorithmes d'optimisation cherchant à minimiser l'énergie présentés dans le 1^{er} chapitre et

$X=0.8$ qui devrait en principe permettre de mieux remplir les liens et donc à priori augmenter l'économie d'énergie en laissant davantage de liens à faible charge.

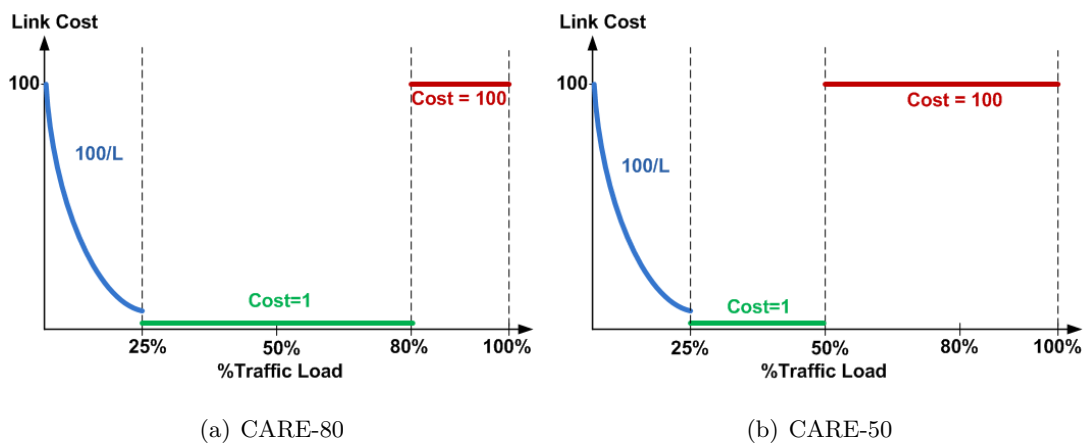


FIGURE 3.3 – Représentation graphique de CARE.

3.1.3 Illustration des politiques d'ingénierie verte proposées

Nous illustrons sur un scénario le routage effectué à l'aide de [OSPF](#) par l'utilisation des métriques [EAGER](#) et [CARE](#).

Nous utilisons une topologie exemple de 3 routeurs qui utilisent des cartes [EEE](#) 1000Base-T.

Le trafic est constitué de 3 flux de débit 160 Mbps dans le scénario a) et de 500Mbps dans le scénario b) pour illustrer respectivement un cas de réseau faiblement chargé (charge de lien inférieure à 25%) et un cas de réseau plus fortement chargé (50% avec un flux).

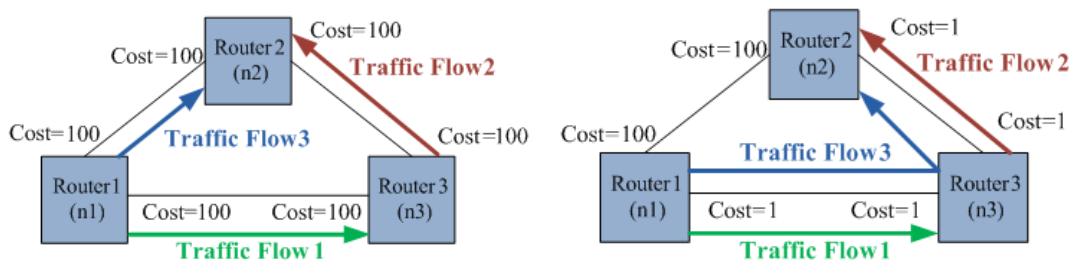
Le premier flux F1 est émis par n1 à destination de n3. Les flux F2 et F3 sont respectivement émis par n3 et n1 à destination de n2.

Le routage des flux 1 et 2 est dans tous les cas équivalent à celui obtenu avec le routage au plus court chemin sans prendre en compte l'énergie, car lorsque les flux sont installés, l'utilisation des liens est nulle. Par contre après leur installation la charge des liens change et le poids qui leur est associé est fonction de la métrique utilisée. Nous obtenons alors des routes différentes pour le flux 3 avec le routage **EAGER** et le routage **CARE**.

- Routage avec métrique **EAGER**

Dans le scénario faible charge a), le trafic suit le plus court chemin en nombre de liens car tous les liens ont un coût équivalent de 100, l'installation des flux F1 et F2 ne modifie pas le coût car la charge des liens reste en deçà de 25% (seuil de modification des poids). Chaque lien étant faiblement chargé peut passer en mode basse consommation, il n'y a donc pas besoin de modifier son poids.

Dans le scénario forte charge b) les flux ont un débit plus important. Lorsque le flux 1 est établi, la charge du lien est de 50%, le coût devient alors 1 : ce lien n'est pas susceptible de se mettre en sommeil alors autant le sélectionner pour faire passer un trafic additionnel. De même, lorsque le flux 2 est établi, le coût du lien entre n3 et n2 passe à 1. Le flux 3 emprunte alors le chemin plus long en nombre de liens (n1-n2-n3) mais qui permet de laisser au repos le lien entre n1 et n2.



(a) 160 Mbps Traffic Flows - Per flow load below the green threshold of 25% (b) 500 Mbps Traffic Flows - Per flow load over the green threshold of 25%

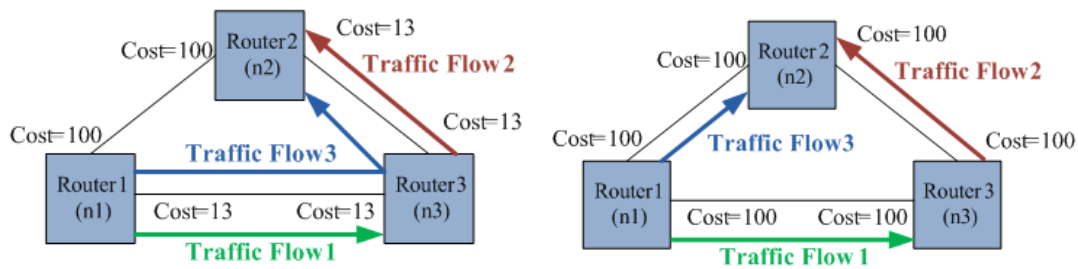
FIGURE 3.4 – Illustration de OSPF-EE avec la métrique **EAGER**

- Routage avec métrique **CARE**

Avec cette métrique, le routage obtenu pour le flux 3 est à l'inverse du routage précédent. A faible charge le flux 3 va passer par le chemin $n1-n2-n3$ au lieu de prendre comme précédemment le chemin direct $n1-n3$ et inversement à forte charge, il empruntera le chemin $n1-n3$. Ceci est dû à la prise en compte de la valeur de la charge dans le calcul du coût et au mécanisme de seuil de congestion que nous avons pris pour l'exemple à 50% ($X=0.5$).

A faible charge les coûts des liens empruntés par F1 et F2 sont modifiés (valeur 13 sur la figure). Ce qui n'était pas le cas avec la métrique précédente. Le routage CARE va tenter de continuer à remplir ce lien car il peut toujours se mettre en sommeil, il sera privilégié dans le calcul de plus court chemin devant un lien moins chargé.

A forte charge, après l'établissement de F1 et F2, les liens sont susceptibles de congestion, leur coût est alors mis à la valeur la plus haute de façon à éviter de passer par ces liens.



(a) 160 Mbps Traffic Flows - Per flow load below the green threshold of 25% (b) 500 Mbps Traffic Flows - Per flow load over the green threshold of 25%

FIGURE 3.5 – Illustration de CARE-50.

Problème et discussion

Les métriques que nous avons définies prennent indirectement en compte la capacité des liens, puisque la charge mesurée sur un lien est fonction de son débit. Un même trafic transmis sur un lien à faible débit induira une charge plus importante sur ce lien que celle qui serait induite sur un lien à débit important. Il n'est cependant pas assuré que ces politiques ne génèrent pas de problème de délai.

Dans l'exemple que nous venons de prendre les débits des flux sont connus, ils ont été choisis pour illustrer les mécanismes d'ingénierie de trafic que nous proposons. Cependant il n'y a pas de certitude que ce fonctionnement soit réellement mis en oeuvre. En effet si le réseau ne met pas en place de mécanisme de contrôle d'admission, permettant de connaître le débit d'un flux (via par exemple l'utilisation d'un seau à jeton), il est possible que le débit du flux soit trop important pour que le résultat de l'ingénierie que nous proposons soit celui escompté. Ainsi, la politique de [CARE](#) est de remplir en priorité jusqu'à un certain niveau (X) les liens en considérant qu'ils vont pouvoir dormir. Hors si le flux arrivant est trop important le lien ne pourra pas dormir, et un délai supplémentaire aura été induit, car le flux traverse un lien supplémentaire. Par ailleurs il faut prendre en compte que le passage par deux liens peut s'avérer plus gourmand en énergie que par un lien.

De façon à mieux appréhender le fonctionnement des métriques et savoir si elles ont le comportement escompté nous avons mené une campagne de simulation permettant de rendre compte de plusieurs choix de routes et plusieurs trafics.

Nous souhaitons évaluer le gain d'énergie, la pénalité en délai et savoir si cette politique obtient des performances similaires à la politique de coalescence que nous avons étudiée dans le chapitre 2. Pour ce faire nous allons effectuer une évaluation de performance en trois étapes. La première étape est une évaluation à partir de modèle de topologie et trafic, la deuxième étape est l'expérimentation sur un banc de test et finalement nous utiliserons un réseau et des trafics réels.

3.2 Evaluation du fonctionnement de l'ingénierie par simulation et expérimentation

Cette section évalue l'ingénierie que nous avons proposée à l'aide de simulation et d'expérimentation sur un scénario similaire à celui que nous avons utilisé pour illustrer le fonctionnement de notre ingénierie verte. L'évaluation par simulation repose en premier lieu sur un modèle de topologie et de trafic. Une évaluation avec des trafics et topologies réels sera présentée par la suite.

Nous utilisons [NS3](#) [[NS3](#)] en tant que plate-forme de simulation dans laquelle nous avons réalisé une mise en oeuvre de la norme IEEE 802.az pour évaluer le fonctionnement de [EAGER](#) et [CARE](#) en termes d'économies d'énergie, délai de bout en bout et répartition de charges.

Dans toutes les expériences, l'économie d'énergie est calculée par rapport à un système qui ne met en oeuvre aucun mécanisme d'économie (Ethernet traditionnel). Ainsi une politique de routage [OSPF](#) de base (que nous appelons legacy) générera un gain en raison du fonctionnement de sa carte Ethernet à économie d'énergie. Notre objectif est de montrer que les politiques proposées obtiennent des gains supérieurs à la politique legacy.

3.2.1 Evaluation par simulation

Pour mener à bien l'étude d'évaluation, nous avons dû faire plusieurs ajouts de code au simulateur. Nous avons ajouté la mise en oeuvre de la norme IEEE 802.3az qui suit la spécification standard [\[eee10\]](#) présentée dans le chapitre 2, une politique de coalescence avec l'algorithme lazy start (voir 1.3.1) et modifié la mise en oeuvre [OSPF](#) pour utiliser les coûts [EAGER](#) et [CARE](#).

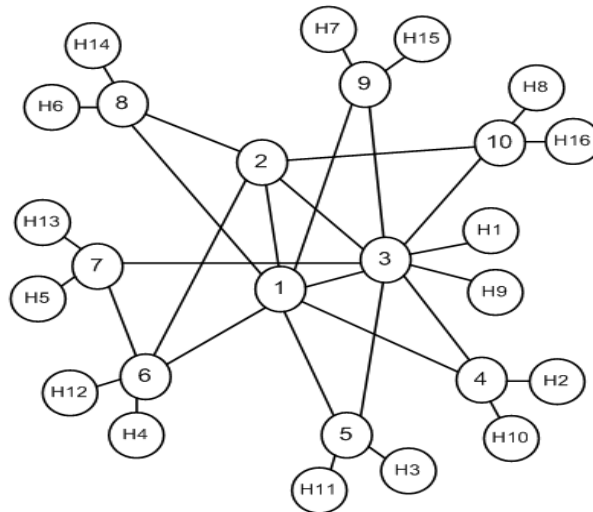


FIGURE 3.6 – Topologie de 10 routeurs générée par BRITE avec le modèle Barabasi-Albert. 16 machinés sont connectées à 8 routeurs de frontière

Topologie

Le simulateur [NS3](#) fonctionne sur une topologie de réseau de 10 routeurs générée à partir du modèle Barabasi-Albert [\[BA99\]](#) par l'outil [Boston university Representative Internet Topology generator](#)

(BRITE) [MLMB01] [BRI] utilisé très largement par la communauté académique pour générer des topologies de réseau internet. Le nombre de routeurs est équivalent à celui du réseau réel ABILENE. L'interface des noeuds est [EEE](#) 1000base-T. La topologie contient 26 noeuds, avec 16 noeuds hôtes (H1-H16) émetteurs et destinataires de flux de trafics et 10 routeurs.

Modèle de trafic

Le trafic est de type [User Datagram Protocol \(UDP\)](#) avec une génération de paquets selon une loi de poisson. 100000 paquets sont transférés avec une taille de 1000 octets.

Le choix de la source est aléatoire parmi les 16 noeuds, 15 choix de source ont été effectués. la source choisie génère un trafic d'entrée sur le réseau variant entre 100 Mbps, 500 Mbps et 700Mbps. Le trafic est décomposé en 7 flux de débit équivalents à destination de tous les destinataires. Ainsi un trafic de 700Mbps correspond à 7 flux de 100 Mbps de la source choisie vers toutes les destinations possibles.

Le choix de la source est effectué de façon cyclique. La durée du temps séparant le choix d'une nouvelle source est estimée en fonction du temps nécessaire pour transmettre le flux de la source. Selon le temps de relayage dans le réseau et les temps d'attentes lors de l'agrégation de trafic sur le lien, la nouvelle source peut commencer à transmettre son trafic alors que les flux de la première source ne sont pas écoulés. Ce qui peut induire des pertes par congestion. Dans le cas pire où tous les flux, 7 flux de 100 Mbps de la source i et 7 flux de 100 Mbps de la source $i+1$, passent par le même lien le trafic est de 1,4 Gbps pour un débit de lien de 1Gbps.

L'intérêt de ce modèle de trafic est de générer des surcharges de trafic dues au multiplexage statistique est de voir comment ces surcharges sont gérées par les politiques d'ingénierie. Selon le débit d'entrée le réseau se retrouve en charge faible, moyenne et haute. Un modèle de trafic par trace réel sera présenté par la suite, avec des charges moins élevées.

Mesure de charge

La méthode de mesure de charge utilisée dans le simulateur est de compter le nombre de bits sur un intervalle de temps. Pour chaque génération de source le temps de transmission est estimé, puis

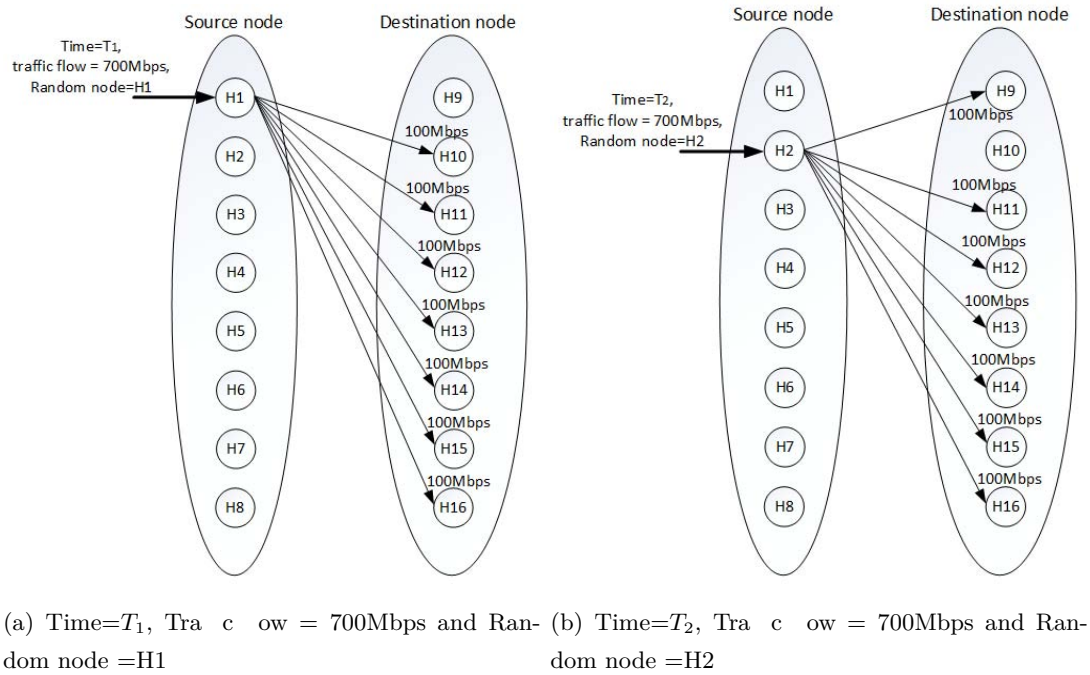


FIGURE 3.7 – Illustration du processus de génération de trafic.

est divisé en 15 intervalles de mesures. Les poids du protocole [OSPF](#) sont calculés à chaque intervalle et transmis par envoi périodique d'annonce.

Notons que pour un trafic d'entrée de 500 Mbps, la charge qui théoriquement devrait être au maximum à la valeur de 50% en début de transmission (une seule source émet), peut être dans le simulateur estimée selon l'instant de mesure à 50,1% et déclencher le changement de poids de 1 à 100 ou bien être de 49,1%. De plus en fin de durée de transmission, le trafic de la $i^{ème}$ source choisie aléatoirement est encore présent dans le réseau quand la $(i + 1)^{ème}$ source commence à émettre, les liens peuvent alors se retrouver chargés à plus de 50% comme l'indique la figure [3.9\(b\)](#).

$$temps\ de\ transmission = \frac{nombre\ de\ bits}{taux\ de\ trafic\ minimum} \quad (3.4)$$

Résultats

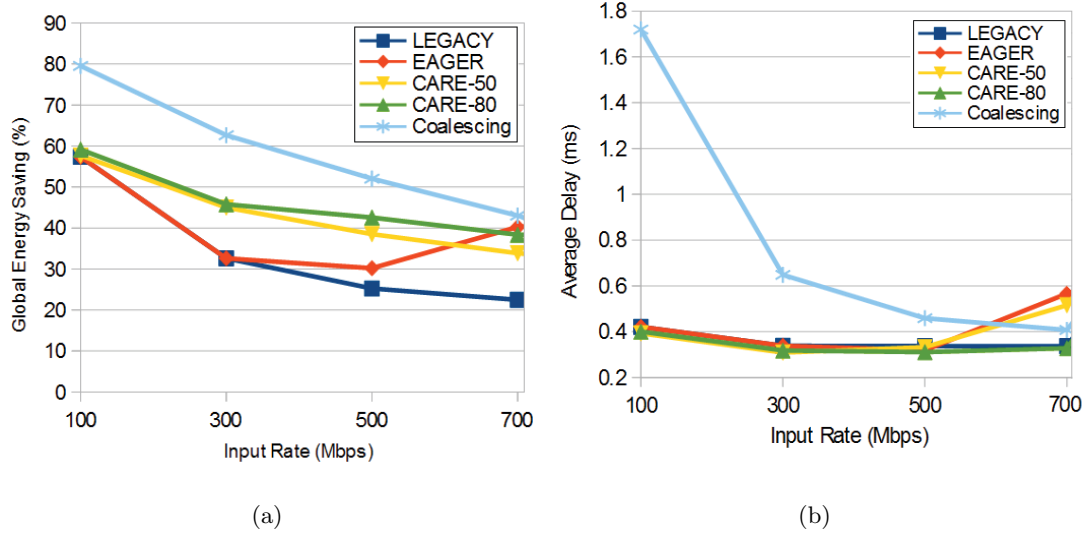


FIGURE 3.8 – Comparaison sur une topologie modèle Barabasi des politiques d'ingénierie verte EAGER et CARE avec le mécanisme de coalescence et avec aucun mécanisme (legacy) - (a) économies d'énergie (b) délai de bout-en-bout moyen

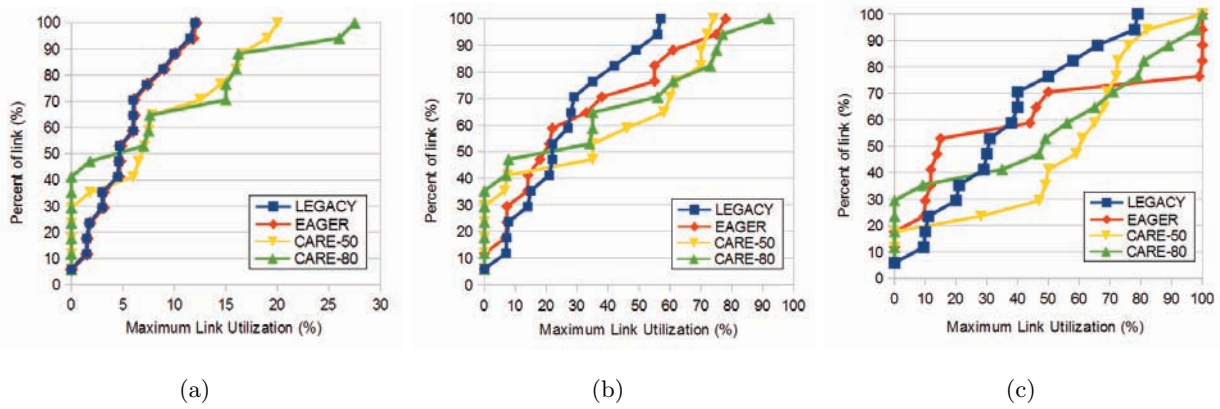


FIGURE 3.9 – Utilisation des liens-Topologie Modèle Barabasi- a) 100Mbps b) 500Mbps c) 700Mbps

Les résultats illustrés en figure 3.8 sont des moyennes obtenues en faisant varier la localisation des sources et destinations et donc des chemins. Ils comparent la performance en terme d'économie d'énergie et de délai moyens de la politique d'ingénierie en version CARE et EAGER avec un fonctionnement sans ingénierie verte, où la politique de routage est au plus court chemin, elle est notée

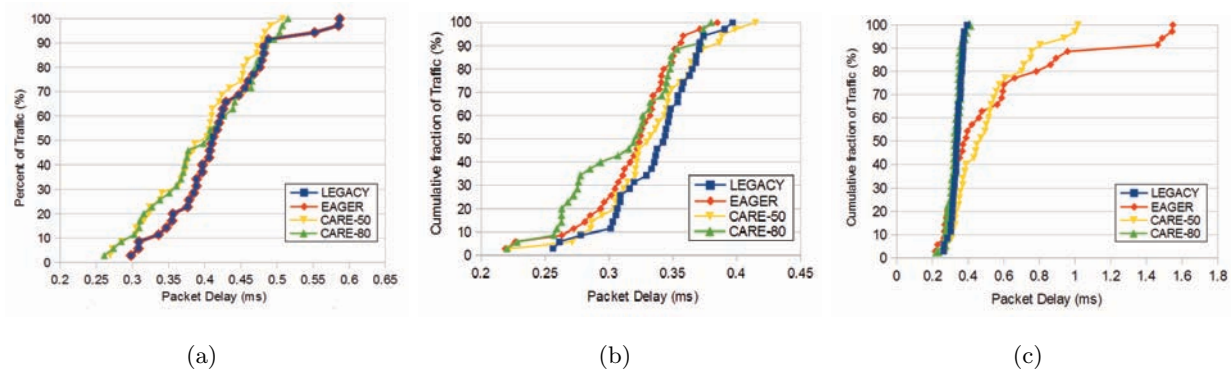


FIGURE 3.10 – Délai de bout en bout- a) 100Mbps b) 500Mbps c) 700Mbps

legacy sur les figures. Toutes les politiques utilisent un Ethernet en mode basse puissance. Nous indiquons également les performances du mécanisme de coalescence vu dans le chapitre 2 où chaque carte Ethernet peut selon les conditions de trafic augmenter la période de sommeil par mémorisation de trames.

La figure 3.9 présente l'utilisation des liens pour des trafics d'entrée de 100Mbps, 500 Mbps et 700 Mbps, les délais de bout en bout pour les mêmes trafics sont indiqués en figure 3.10. Ces résultats nous permettent de mettre en regard le gain d'énergie d'une politique et son impact sur la qualité de service du réseau.

Les résultats sont analysés selon trois angles :

- Intérêt de mettre en place une ingénierie verte au niveau 3 par rapport à un mécanisme de coalescence de niveau 2.
- Choix de la métrique à utiliser.
- L'impact sur la qualité de service.

3.2.1.1 Ingénierie verte de niveau 3 versus Coalescing de niveau 2 A faible charge d'entrée (100Mbps en entrée décomposé en 7 flux de 100/7 Mbps), les liens sont au maximum utilisés à 10% avec un routage classique (cf Fig 3.9 a), nous notons que l'ingénierie verte n'apporte pas de gain par rapport au routage au plus court chemin, le gain obtenu (global energy saving) noté en

ordonné provient du mécanisme d'endormissement Ethernet de niveau 2 qui est présent dans toutes les politiques évaluées.

Dès que la charge augmente la politique d'ingénierie verte devient intéressante, son gain étant de 1.5 à 2 fois plus important qu'avec une politique au plus court chemin.

Comparé au mécanisme de coalescence, le gain est nettement moins important à faible charge. Le gain de la coalescence diminue avec la croissance de la charge du réseau mais subsiste.

Il apparaîtrait donc que pour améliorer l'économie d'énergie il vaut mieux utiliser un mécanisme au niveau 2 qu'une politique de niveau 3. Cependant la performance en délai (Fig 3.10) du mécanisme de coalescence est nettement moins bonne que celle de l'ingénierie verte, avec une différence très importante à faible charge.

En coalescence, les trames doivent à chaque routeur, attendre avant d'être transmises que le tampon soit rempli jusqu'au seuil, ou que le timer expire. A forte charge la coalescence de niveau 2 ajoute un délai moins important car le trafic est suffisamment important pour que l'attente avant émission soit faible, alors que l'ingénierie verte a un délai légèrement plus important que le fonctionnement de base, car les paquets ne suivent pas le plus court chemin.

3.2.1.2 Politique EAGER versus CARE Pour de faibles charges (trafic entrant inférieur à 300 Mbps) la politique **EAGER** a le même comportement que celle de la politique de base, les poids des liens ne sont pas modifiés, le routage est identique, c'est un routage au plus court chemin.

En charge moyenne les poids sont modifiés pour obtenir un routage différent qui génère alors un gain pouvant aller de 1,5 à 2 fois le gain du routage classique.

Pour une forte charge (de 500 à 700) le gain de la politique **EAGER** augmente. Ceci est conforme au résultat de la figure 3.9 qui indique un nombre de liens à faible charge plus important avec **EAGER** qu'avec les politiques **CARE**. La politique Eager va concentrer le trafic sur les liens chargés au-delà de 25%, alors que les politiques **CARE** au-delà de leur seuil vont répartir le trafic et utiliser davantage de liens tel qu'illustré sur la figure 3.10.

La politique **CARE** est davantage sensible à la charge et modifie les poids des liens selon sa valeur,

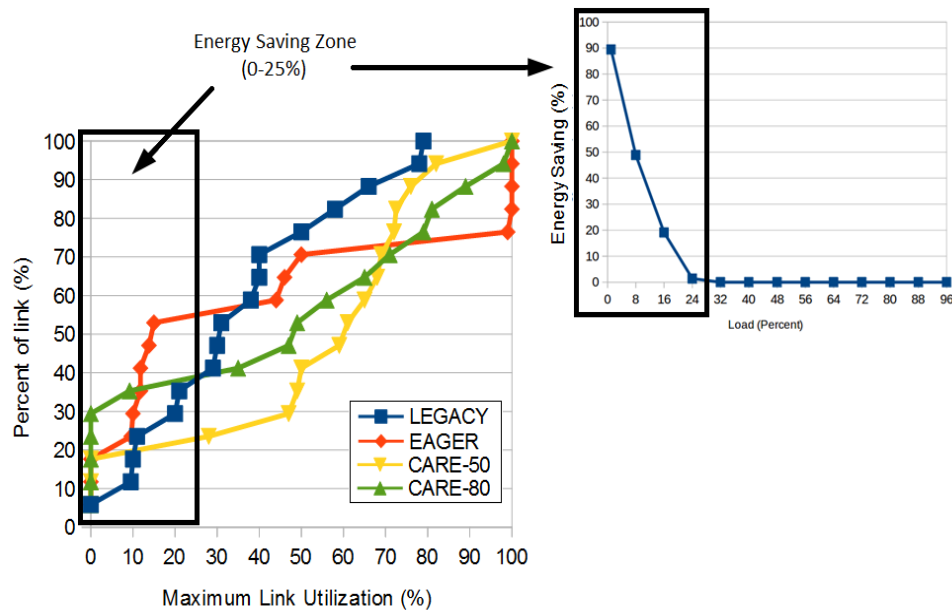


FIGURE 3.11 – Zone d'économie d'énergie-Cas d'un trafic d'entrée à 700 Mbps

générant un gain dès les faibles charges alors que **EAGER** ne modifie pas les poids et suit le routage traditionnel.

3.2.1.3 Ingénierie de trafic et qualité de service

Examinons l'impact des politiques sur la qualité de service.

- Congestion

La charge des liens avec toutes les politiques ne dépasse pas 100%, il n'y a pas de congestion.

- Délai

L'impact sur le délai est lié à la charge. A faible charge, les délais des politiques sont sensiblement équivalents. Pour une charge moyenne une différence apparaît quant au nombre de paquets ayant un faible délai, mais concernant le délai le plus important, le résultat des politiques est similaire.

La différence d'impact des politiques apparait à forte charge. La politique **EAGER** qui n'a pas de seuil d'évitement de congestion est celle qui à forte charge occasionne les délais les plus importants, par contre, la politique **CARE** 80 obtient les mêmes résultats que la politique Legacy. La politique **CARE** 80 augmente le délai (mais moins que la politique **EAGER**).

3.2.1.4 Synthèse des résultats des politiques vertes Nous avons synthétisé l'analyse des résultats en table 3.1 en indiquant des niveaux 'faible' 'moyenne' 'grande' dont les valeurs sont fournies en table 3.2

TABLE 3.1 – Synthèse des performances par niveaux en énergie et délais des métriques

métriques	Economie d'énergie globale			Délai de bout en bout		
	Charge Faible	Charge Moyenne	Charge Haute	Charge Faible	Charge Moyenne	Charge Haute
EAGER	grande	faible	grande	faible	faible	faible
CARE-50	grande	moyenne	moyenne	faible	faible	faible
CARE-80	grande	grande	grande	faible	faible	faible
Coalescing	très grande	grande	grande	très grand	faible	faible
LEGACY OSPF	grande	faible	faible	faible	faible	faible

TABLE 3.2 – Valeurs des niveaux d'économie d'énergie globale et de délai de bout en bout

Economie d'énergie Globale ("G" en %)	Délai de bout en bout ("E" en ms)	Niveaux
$G > 60\%$	$E > 1.275$	très grande
$40\% < G \leq 60\%$	$0.85 < E \leq 1.275$	grande
$30\% < G \leq 40\%$	$0.6735 < E \leq 0.85$	moyenne
$G \leq 30\%$	$E \leq 0.6735$	faible

Conclusion

Le scénario de trafic simulé correspond à un fonctionnement de charge faible, moyenne, et forte qui donne les résultats escomptés selon la définition des métriques.

A très faible charge, il n'y a pas besoin de mettre en place de politique verte, l'endormissement des fonctions d'émission et réception permet d'économiser de l'énergie. Les politiques sont avantageuses à charge moyenne et forte avec plus ou moins d'impact sur la qualité de services.

La méthode **EAGER** simple à mettre en oeuvre est adaptée à une charge de trafic moyenne, mais occasionne une augmentation de délai significative lorsque la charge augmente.

A forte charge la politique **CARE** est préférable. Le compromis entre augmentation de délai et gain d'énergie est représenté par le choix de la métrique. Plus celui-ci est élevé plus l'économie est importante mais plus le délai risque d'être accru. Sur ces simulations la solution **CARE** 80 semble être un bon compromis.

Lorsque la charge augmente l'ingénierie verte permet quasiment de doubler le gain obtenu par le standard IEEE 802.az

Nous complétons à la suite notre évaluation par des mesures réelles.

3.2.2 Evaluation expérimentale

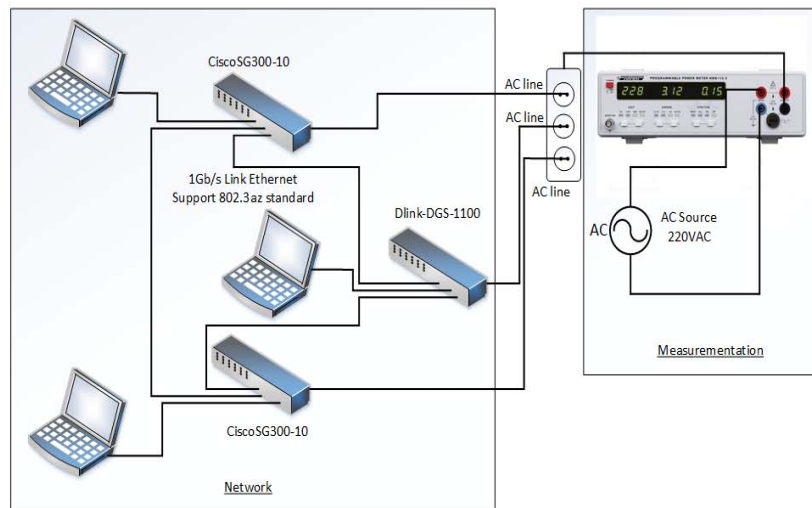


FIGURE 3.12 – Eléments du banc de test

En raison du matériel disponible nous avons effectué l'expérimentation à partir de deux commutateur de niveau 3 (i.p., Cisco SG300-10) et d'un commutateur de niveau 2 (à savoir, Dlink

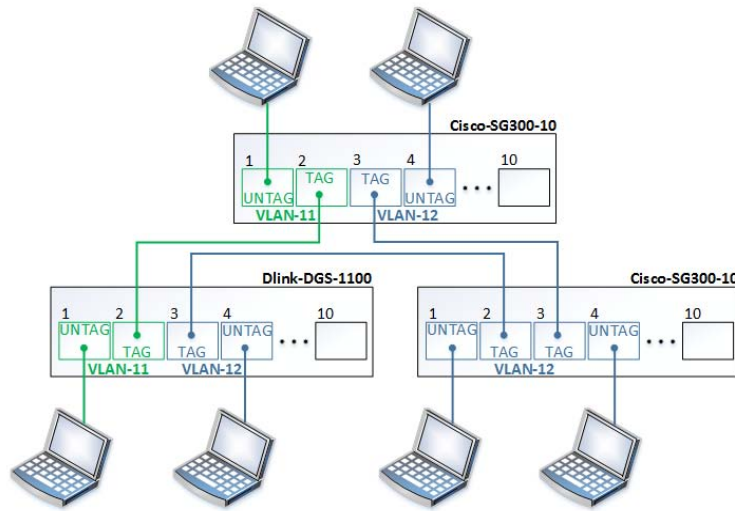


FIGURE 3.13 – Configuration du banc de test

DGS-1100) comme le montre la figure 3.13. 5 tests ont été effectués et moyennés, (les fluctuations de consommation ne sont pas significatives).

Expérimentations et résultats

L'objectif est d'observer le comportement énergétique en comparaison avec le fonctionnement du routage **OSPF** LEGACY. Pour ce faire nous avons repris le scénario qui nous a servi à illustrer les politiques **EAGER** et **CARE** en section 1.1.3. et effectué des configurations de routage en utilisant des **VLAN**, selon les figures 3.14 et 3.15.

Nous avons mis en oeuvre 4 tests dont les résultats sont indiqués en Figure 3.16.

- Test 1 : pour examiner le comportement **OSPF** en générant des flux inter **VLAN** (cf figure) avec un routage au plus court chemin.
- Test 2 : pour observer la politique **EAGER**,
charge $\leq 25\%$: nous utilisons la configuration de LEGACY **OSPF** (Fig.3.14)
charge $> 25\%$: la configuration est illustrée en Figure 3.15

- Test 3 : pour observer **CARE** 50,
Charge $\leq 50\%$: figure 3.15.
Charge $> 50\%$: figure 3.14.
- Test 4 : observation de **CARE** 80,
Charge $\leq 80\%$: nous utilisons la configuration pour **CARE** (Fig.3.15).
Charge $> 80\%$: nous utilisons la configuration **LEGACY OSPF** (Fig.3.14).

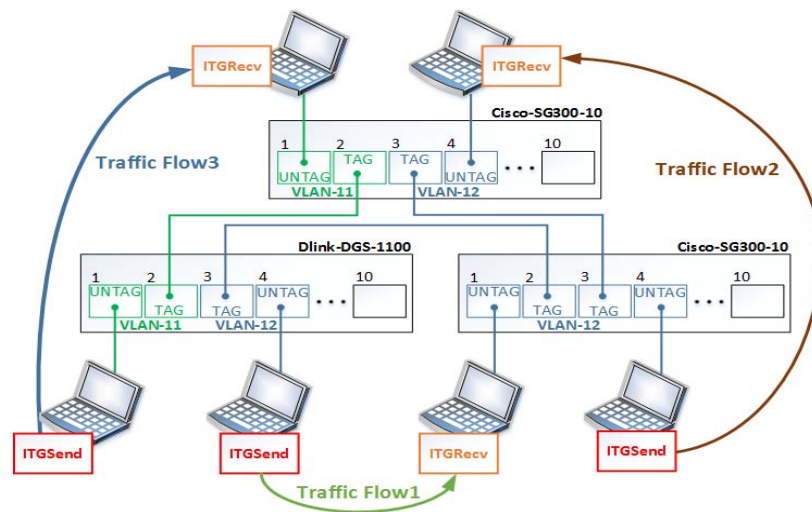


FIGURE 3.14 – Configuration pour LEGACY OSPF

Les résultats d'expérimentation sont illustrés en figure 3.16

Nous constatons que le mécanisme IEEE 802.3az peut réduire considérablement la consommation d'énergie, même dans une topologie simple. Cependant, cet avantage disparaît avec l'augmentation de charge qui réduit les opportunités pour les cartes de ligne de passer en mode de faible puissance.

EAGER et **CARE**, en acheminant le trafic sur des liaisons qui sont déjà utilisés (via le routeur 3 à la place du lien direct) doublent pratiquement le pourcentage d'énergie économisée.

Quand le trafic augmente, comme prévu, en premier **CARE**-50 puis **CARE**-80, fonctionnent comme **OSPF** LEGACY. **EAGER** par ailleurs ne modifie pas son routage lorsque le lien est très

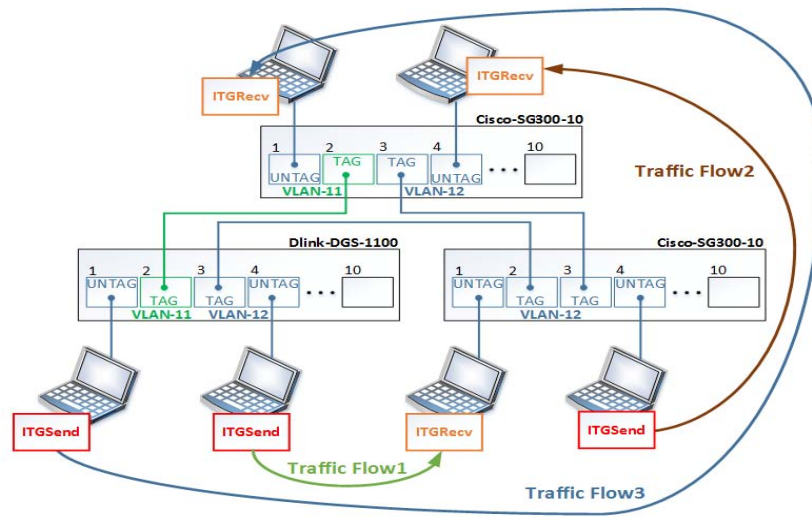


FIGURE 3.15 – Configuration pour EAGER et CARE

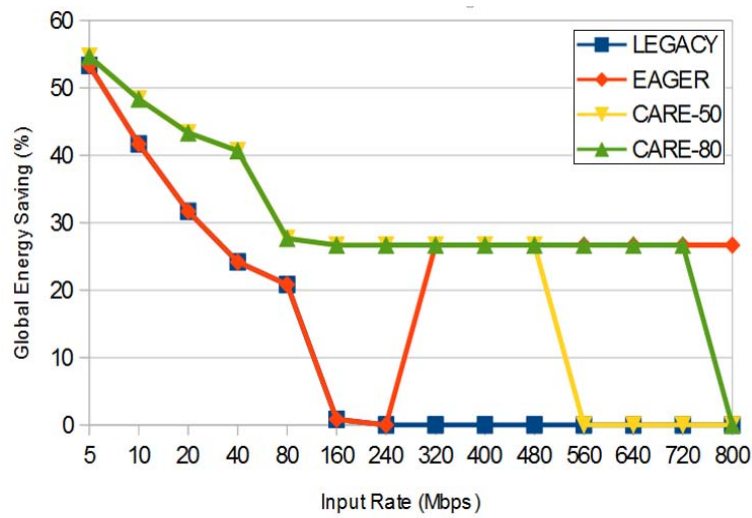


FIGURE 3.16 – Pourcentage d'économie d'énergie globale du réseau de test.

encombré (contrairement aux politiques CARE) ce qui donne l'occasion aux deux cartes de N1 à N2 de passer en mode LPI, augmentant ainsi l'économie d'énergie à forte charge.

Conclusion : Le fonctionnement de notre politique est évalué comme correct, nous allons à présent chercher à savoir si les résultats qui sont obtenus sont loin d'une solution optimale, et s'ils apportent

quelque chose par rapport à une solution existante.

3.3 Comparaison de la proposition sur un scénario réaliste

Nous avons analysé le fonctionnement de nos propositions avec des conditions de charge et un modèle de topologie qui nous as permis de mener à bien notre étude mais qui ne sont pas représentatifs de l'utilisation actuelle des réseaux de coeur. Nous allons à présent confronté nos travaux à des conditions réelles et en préciser l'intérêt.

Dans cette section, nous considérons un scénario réaliste de topologie et de trafic en se référant au réseau Abilene (Internet 2). *Nous confrontons notre politique d'économie d'énergie à 1) une politique d'ingénierie verte idéale et 2) à une politique issue de la littérature, Green TE, obtenue par optimisation.*

3.3.1 Comparaison à une solution de mesure idéale

Nous allons comparer les résultats des modèles d'économie d'énergie [EEE](#) que sont [EAGER](#) et [CARE](#) avec une modélisation idéale, que nous présentons à la suite. Nous considérons cette métrique comme étant idéale pour une politique d'ingénierie verte [IP](#) en réseaux [EEE](#) dans la mesure où la consommation d'énergie correspond au plus près de la réalité.

3.3.1.1 Modélisation idéale de l'économie d'énergie La modélisation que nous qualifions d'idéale consiste à être capable de déterminer finement la charge à tout instant et en déduire l'économie d'énergie. En se référant à la figure [3.1](#) nous utilisons le développement en série de Fourier fourni par MathLab ci-dessous :

$$f(x) = a_0 + \sum_{m=1}^{\infty} [a_m \cos(wx) + b_m \sin(wx)] \quad (3.5)$$

où $f(x)$ est l'économie d'énergie obtenue pour une charge x . Elle est égale à la somme des quatre premiers termes de la série de Fourier et a_m et b_m sont les coefficients de Fourier de la série.

Nous définissons la métrique IDEAL par

$$IDEAL_{i,j} = \begin{cases} 0 + \sum_{m=1}^{\infty} [{}_m \cos(wx) + {}_m \sin(wx)] \\ 1 \quad \text{if } IDEAL_{i,j} < 1 \end{cases} \quad (3.6)$$

Dans nos simulations nous allons comparer les performances obtenues avec les modèles [EAGER](#) et [CARE](#) à celles obtenues avec cette métrique que nous pouvons calculer par simulation, mais qui se prête difficilement à une implantation réelle (trop complexe).

3.3.1.2 Paramètres de Simulation La topologie et les trafics simulés sont les suivants.

Réseau Abilene

Le réseau simulé avec [NS3](#) est le réseau Abilene (Fig 3.17) situé en Amérique du nord qui a servi de base à plusieurs études [[ZYLZ10](#)] [[SLX⁺12](#)] [[RRCRS14](#)] [[Yin](#)]. Il est constitué de liens 10 Gbps reliant 11 routeurs (12 routeurs selon certains qui séparent 2 routeurs qui sont co-localisés) que nous supposons équipés de cartes [EEE](#).

La moyenne d'utilisation des liens est autour de 2% avec un maximum qui fluctue dans la plupart des cas entre 10% et 20% et exceptionnellement des pointes au delà de 50%. Ce type de comportement est représentatif dans les réseaux commerciaux.

La redondance des chemins et leur faible taux d'occupation font d'Abilene un bon candidat à de l'ingénierie verte. Par ailleurs, ce choix de réseau nous permet une comparaison de nos résultats à ceux obtenus en [[LSLC13](#)].

Matrice de trafic

La matrice de trafic estimée ([Traffic Matrix \(TM\)](#)) de la topologie du réseau Abilene couvre une période de 24h, elle reflète le trafic des onze sources vers les dix destinations. Nous avons

- **TM** : Une matrice de $N \times N$ nombres entiers non-négatifs, avec $T[i, j]$ le nombre de messages envoyés à partir du noeud i au noeud j dans la période d'observation. Nous avons généré la matrice selon la méthode proposée en [[Yin](#)].

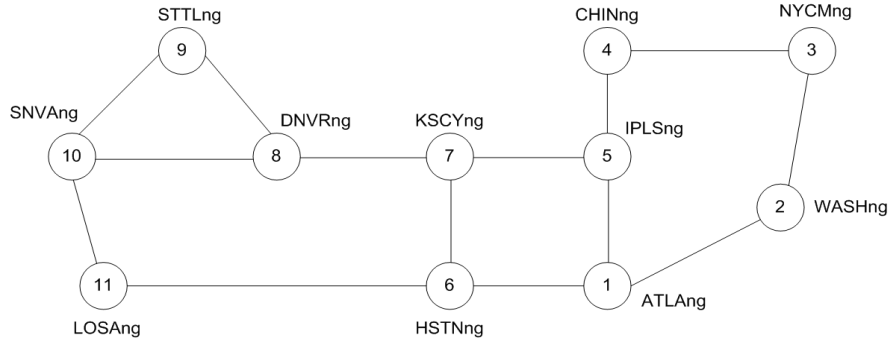


FIGURE 3.17 – Topologie du réseau Abilene

- **Routes** : Si la matrice de trafic indique que $T[i, j]$ messages doivent être envoyés à partir du noeud i au noeud j dans une période de temps, ces messages doivent être acheminés à partir du noeud i au noeud j . La route du noeud i au noeud j est un chemin dans le graphe de la topologie du réseau j . Une affectation de débit est spécifiée pour chaque chemin.
- **Charge du lien** : La charge sur un lien (simplex) est la somme du nombre de messages délivrés sur les chemins qui comprennent ce lien. La charge sur un lien ne doit pas dépasser sa capacité.
- **Charge totale** la charge totale de trafic dans une matrice de trafic $N \times N$ T est

$$L(T) = \sum_{i,j \in [1, N]} T[i, j] \quad (3.7)$$

La charge totale (ou agrégée) est simplement la somme des charges de liaison.

3.3.1.3 Intérêt de la modélisation CARE 80 Les résultats de simulation indiqués en Fig.3.18, montrent une quasi coïncidence entre d'une part le routage IDEAL et le routage CARE 80, et d'autre part, le routage EAGER et le routage LEGACY. Néanmoins, avec un maximum d'utilisation de lien, (MLU) supérieur à 20%, le pourcentage d'économie d'énergie réalisée avec EAGER est plus élevé que celui avec le routage LEGACY.

L'économie d'énergie moyenne obtenue via le modèle CARE-80 et le modèle CARE-50 est environ

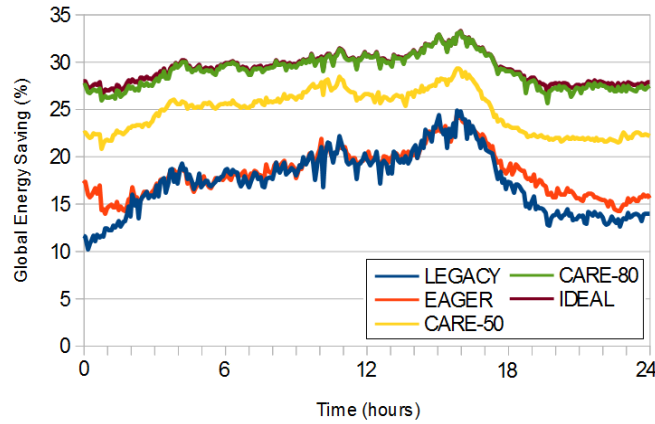


FIGURE 3.18 – Economie d'énergie globale sur le réseau Abilene.

2,5 fois plus importante que celle obtenue par les deux autres routages lorsque la charge du lien est importante (représentée en figure 3.19).

L'économie d'énergie maximale atteinte sur le réseau Abilene avec la fonction métrique IDEAL est de 34% pour des *MLU* faibles ($\sim 14\%$). Lorsque le *MLU* augmente au-delà de 25%, l'économie d'énergie chute à 26%.

Concernant l'occupation des liens (Figure 3.19, Fig 3.20), le réseau est faiblement chargé la plupart du temps (environ 22 sur 24 heures) et l'utilisation maximale de liaison, en moyenne, est inférieure à 20%. En outre, l'utilisation maximale de liaison pour CARE-80 et IDEAL culmine à près de 70%.

A partir de ces résultats nous interprétons le fonctionnement des politiques par :

La politique legacy dirige le trafic sur le plus court chemin. Lorsque les plus courts chemins entre source et destination sont disjoints ou partiellement disjoints de n , la charge est répartie entre n chemins.

La politique CARE privilégie certains des chemins parmi les plus courts chemins, en y concentrant le trafic.

La politique EAGER répartit la charge sur plus de chemins que la politique Legacy en utilisant, d'autres chemins que les plus courts chemins car la charge dépasse par moment le seuil des 25%(cf

figure 3.20). Selon les heures creuses ou pleines respectivement les liens sont chargés à moins 60% et à moins 40%.

Conclusion : Sur le cas réel de Abilene, la politique CARE est proche du comportement avec une métrique idéale. Elle tente de charger les liens et est plus intéressante que la politique EAGER qui répartit le trafic sur des liens peu chargés. Ses performances sont quasi idéales. Nous relevons que aucune politique ne crée de congestion.

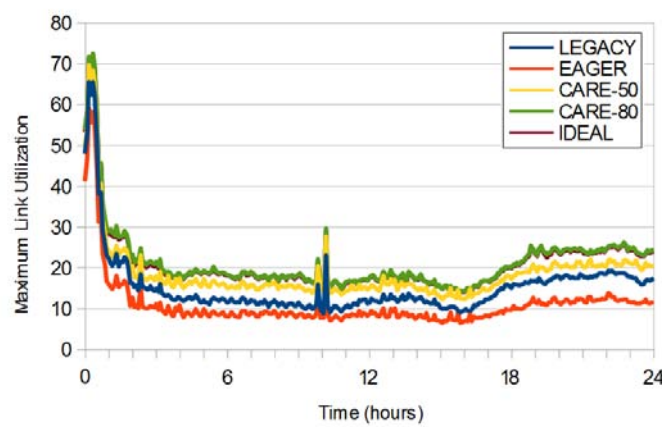


FIGURE 3.19 – Maximum link utilization du réseau Abilene mesuré toutes les 5 minutes

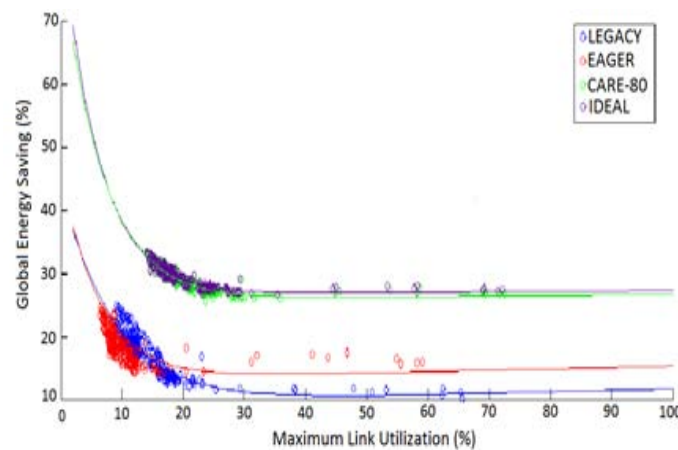
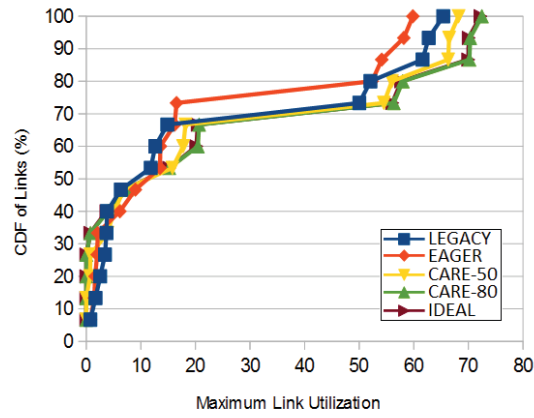
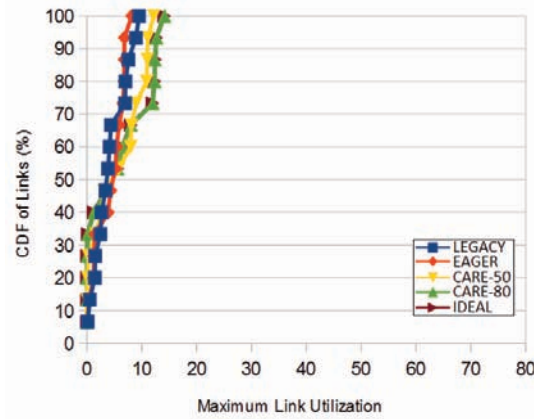


FIGURE 3.20 – Compromis entre MLU et économie d'énergie pour la topologie du réseau Abilene



(a) Heures pleines



(b) Heures creuses

FIGURE 3.21 – Cumulative distribution functions (CDF) de l'utilisation maximale des liens a) en heures creuses (b) en heures pleines

3.3.2 Comparaison à la solution optimale Green-TE

Nous comparons à présent l'intérêt d'effectuer une ingénierie dépendante de la charge et de la technologie [EEE](#) par rapport à l'utilisation d'une ingénierie ON/OFF indépendante de la technologie mais optimisée.

Dans [\[ZYLZ10\]](#), Zhang propose d'optimiser l'énergie en maximisant le nombre de liens qui peuvent être mis en sommeil sous des contraintes de performances d'utilisation de la liaison ([MLU](#)) et le délai.

L'objectif est d'éteindre les cartes d'interfaces. Ces cartes gèrent plusieurs liens (cas d'aggrégation de liens).

Le problème d'optimisation est formulé sous forme [Mixed Integer Programming \(MIP\)](#) comme suit :

$$\text{Minimize} \quad \sum_{l \in E} P_l x_l + \sum_{m \in M} B_m y_m \quad (3.8)$$

$$\text{S.t.} \quad \sum_{l \in E} f_l^{s,t} O_l^i - \sum_{l \in E} f_l^{s,t} I_l^i = \begin{cases} D_{s,t} & \text{if } i = t \\ -D_{s,t} & \text{if } i = s \\ 0 & \text{if } i = s, t \end{cases} \quad (3.9)$$

$$\sum_{l \in S_m} y_m \leq x_l \quad (3.10)$$

$$u_l = \frac{1}{C_l} \sum_{s,t \in V, s \neq t} f_l^{s,t} \quad , \quad l \in E \quad (3.11)$$

$$x_l = x_{r(l)} \quad , \quad l \in E \quad (3.12)$$

$$x_l + u_l \leq 1 \quad , \quad l \in E \quad (3.13)$$

3-8 : c'est l'objectif d'économie d'énergie dans le réseau, le modèle d'énergie rend compte de cartes multiports ; les résultats sont obtenus avec des cartes monoports

3-9 : indique les contraintes de conservation de flux.

3-10 : $|S_m|$ est le cardinal de S_m , l'équation assure qu'une carte gérant plusieurs liens est mise en sommeil que lorsque tous ses ports sont endormis. Les résultats obtenus par les auteurs l'ont été avec une valeur S_m de 1. 3-11 calcule l'utilisation de la liaison. 3-12 s'assure que les liaisons sont mises en sommeil par paire. Il n'y a pas de circulation de trafic ni entrant ni sortant.

3-13 : stipule qu'un lien peut être mis en sommeil que s'il n'y a pas de trafic et que quand le lien est actif, le trafic qu'il supporte est inférieur à sa capacité.

Le modèle général considère tous les chemins possibles pour chaque paire origine-destination, mais afin de réduire la complexité de l'optimisation les auteurs pré-calculent un ensemble de chemins candidats et recherchent des solutions au sein de cet ensemble. Trois heuristiques de choix de chemins candidats sont proposées par les auteurs.

- **basic** : la recherche est effectuée parmi K plus courts chemins
- **basic + nd** : la recherche est effectuée parmi K plus courts chemins qui satisfont une contrainte de diamètre de réseau (nd).
- **basic + e2e** : la recherche est effectuée parmi K plus courts chemins qui sont également conformes aux contraintes de délai (e2e).

Les performances de ces 3 heuristiques sont évaluées sur le réseau Abilene et indiquées en Figure 3.22 avec un MLU de 50% pour la contrainte d'optimisation.

Les auteurs indiquent que leurs heuristiques basic et basic + nd sont calculées avec une valeur de K suffisamment grande pour que les résultats soient optimaux, c'est-à-dire équivalents à ceux obtenus par résolution du problème avec l'outil CPLEX (noté MCF-TE en Figure 3.23). GreenTE est capable d'atteindre environ 27% d'économie d'énergie. Par contre l'heuristique s'appuyant sur le délai ne sélectionne pas suffisamment de candidats et obtient un résultat environ 20%, inférieure aux deux autres.

Les auteurs de GreenTE calculent l'économie d'énergie en comparant la puissance totale c'est-à-dire celle consommée par toutes les cartes lorsque celles-ci sont en sommeil en regard de la puissance consommée par toutes les lignes en activité. Les valeurs sont des valeurs de puissances indiquées pour des routeurs CISCO, pour notre part nous utilisons les valeurs de puissance Ethernet et calculons l'énergie. Cependant dans les deux cas, sont exprimés des rapports, qui peuvent être comparés. Nous indiquons les résultats que nous obtenons dans les Figures 3.22 et 3.23.

Conclusion : La politique CARE-50 est la politique la plus proche de la politique Green-TE en ce sens que les deux politiques ont pour but d'avoir une charge de lien inférieure à 50%. La politique optimale Green-TE obtient des résultats légèrement plus intéressants et la politique CARE 80 obtient une économie légèrement meilleure que la solution optimale Green TE reposant sur un

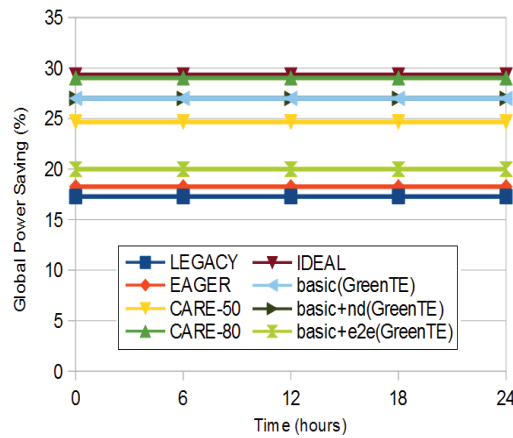


FIGURE 3.22 – Comparaison de performances Green TE, EAGER, CARE

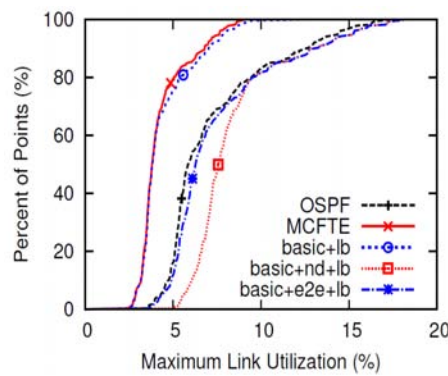


FIGURE 3.23 – Utilisation des liens

modèle ON/OFF.

La politique ON/OFF considère que la liaison ne consomme pas d'énergie lorsqu'elle est inutilisée, les équipements sont désactivés. Cette désactivation pose des problèmes de gestion de la robustesse, car il faut pouvoir détecter la panne d'un lien, et réactiver rapidement un lien stoppé pour qu'il puisse servir de secours. Il nous semble préférable d'utiliser le mécanisme d'endormissement. La consommation du lien n'est pas nulle mais les résultats obtenus par ingénierie sont proches de ceux de la solution ON/OFF.

Notons également que la réalisation de la méthode optimale nécessite une connaissance préalable

du trafic difficile à obtenir dans les réseaux internet, alors que l'ingénierie verte s'appuie sur des mesures de charge.

3.4 Architecture multicouche - Prise en compte de l'énergie lien

Nous avons développé et analysé une métrique d'ingénierie paramétrée sur la charge du réseau et montré qu'elle obtient des résultats proches de Green-TE.

Cependant la métrique ne prend en compte que l'énergie consommée par l'émission et la réception sur le lien, alors que dans la technologie optique déployée dans les réseaux de coeur, le transfert sur le lien est également consommateur d'énergie.

La transmission sur un câble optique suppose l'utilisation d'amplificateurs régénérateurs pour corriger les effets d'atténuation et de dispersion. Le nombre d'amplificateurs nécessaires est fonction de la distance à parcourir. Dans ce chapitre nous considérons qu'un amplificateur est positionné tous les 80Km pour assurer un niveau de signal suffisant à la transmission, [VHIV⁺12] comme indiqué en figure 3.24.

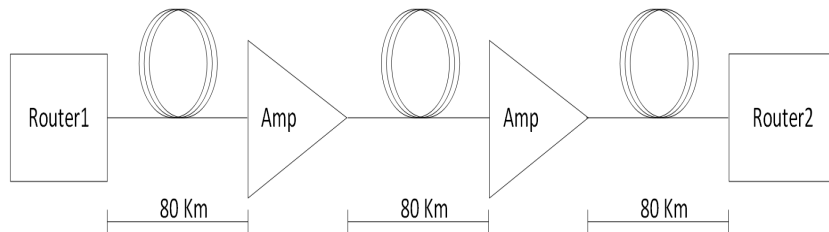


FIGURE 3.24 – Lien en support optique

3.4.1 Proposition de métrique verte en fibre optique

L'idée principale de l'ingénierie optique verte est de pénaliser des liens ayant un grand nombre d'amplificateurs même si ils sont faiblement chargés car leur consommation sera trop importante. Nous continuons également à éviter les liens trop chargés (au-delà de X%) qui ne pourront pas se mettre en sommeil.

Nous proposons de prendre en compte à la fois la consommation de la liaison et des amplificateurs optiques par la formulation de métrique ci-après :

$$Optic \quad CARE \quad X = \begin{cases} \frac{100}{L_{i,j}} \quad AF_{i,j} & , \text{if } L_{i,j} \leq X/2 \\ AF_{i,j} & , \text{if } X/2 < L_{i,j} \leq X \\ 100 \quad AF_{i,j} & , \text{if } L_{i,j} > X \end{cases} \quad (3.14)$$

Avec

- i) $L_{i,j}$ la charge du lien i, j pour prendre en compte le coût de liaison ;
- ii) $AF_{i,j}$ le facteur de coût relatif aux amplificateurs du lien $L_{i,j}$;
- iii) X est le seuil d'utilisation du lien. Nous prenons $X = 80$ et $X = 50$
- iv) les éléments de valeurs sont normalisés

Le coût d'un lien est fonction de sa charge et du nombre d'amplificateur qui le compose, ceux-ci générant un facteur de coût noté **Amplifier Factor (AF)**. Nous faisons varier la valeur de **AF** entre 1 et 10.

- Si la charge du lien est importante, au-delà du seuil de congestion X , quel que soit le nombre d'amplificateur, le lien doit être évité. Son coût sera au minimum de 100 (si il n'y a pas d'amplificateur sur la liaison **AF** a comme valeur 1)
- Si la charge du lien est au-delà de 25%, nous avons constaté que le gain d'énergie par la mise en sommeil ne compense pas l'énergie consommée par le transfert sur un lien plus long avec plus d'amplificateurs, c'est pourquoi nous considérons uniquement dans le coût le nombre d'amplificateur.
- A faible charge nous considérons à la fois le nombre d'amplificateurs et la charge.

Précision sur le facteur amplificateur

Lorsque le nombre d'amplificateur est faible, c'est-à-dire inférieur au tiers du nombre maximal d'amplificateur, le coût induit est faible (**AF** égal à 1). Lorsque le nombre d'amplificateur est important, supérieur au deux tiers du nombre maximal d'amplificateurs, le coût en énergie est prépondérant, la

valeur est maximale (AF égal 10). Pour un nombre d'amplificateur moyen, compris entre un tiers et deux tiers du nombre maximal d'amplificateurs, le coût est exponentiel, il varie de 1 à 5. L'algorithme de calcul de métrique est indiqué ci-contre (algorithme 1).

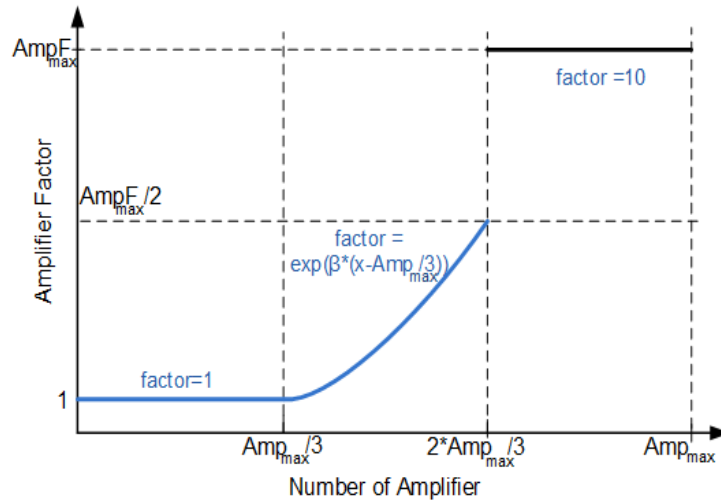


FIGURE 3.25 – Illustration du facteur Amplificateur (AF)

3.4.2 Evaluation

Evaluons à présent cette politique sur le réseau Abilene. Nous cherchons à déterminer sur le réseau réel, le bien-fondé de la prise en compte de la consommation du niveau physique, nous allons comparer notre proposition à 1) notre proposition précédente qui prenait en compte uniquement la charge de trafic dans le calcul de consommation et, à 2) une solution qui ne prendrait en compte que le nombre d'amplificateurs (Seul Ampli).

Résultats et analyse

Nous considérons le réseau conforme à une architecture composée des couches :

- IP
- Ethernet
- Optical Transport Network (OTN)

Algorithm 1 : Allocation verte de coûts en réseau optique

```

1 Input :  $A_i, A_t, L_i, X$  (Number of Amplifier, Amplifier threshold, link utilization, congestion
   threshold)
2 for  $i \in \{\text{link interfaces of current router}\}$  do
3   Search  $A_i$ 
4   link_List.add(link i,  $A_i$ )
5   if  $A_i > \text{Amplifier threshold}$  then
6     | link_metric_i = max value
7   end
8   if  $A_i \leq \text{Amplifier threshold}$  then
9     | if  $L_i$  is low value  $L_i$  is middle value then
10      | Estimate AF_i and LF_i
11      | Compute_link_metric_i =  $AF_i \cdot LF_i$ 
12    end
13    if  $L_i$  is high value then
14      | link_metric_i = max value
15    end
16  end
17 end

```

– Wavelength Division Multiplexing (WDM)

La puissance globale (P_{core}) est calculée en utilisant l'équation ci-dessous avec les valeurs de la table 3.3 [VHIV⁺12].

$$P_{core} = P_{ip} + P_{ethernet} + P_{otn} + P_{wdm} \quad (3.15)$$

$$P_{wdm} = P_{optsw} + P_{transponders} + P_{amplifiers} + P_{regenerator} \quad (3.16)$$

Nous évaluons la consommation d'énergie sur la topologie Abilene en regardant une heure de trafic. Les amplificateurs sont espacés tous les 80km, la distance des liens est indiquée en figure 3.26.

TABLE 3.3 – Consommation d'énergie pour équipements en réseau optique multicouches.

Niveaux de l'architecture	Type	Puissance
Niveau IP	CRS-3 10G par card	5W par Gbps
Niveau Ethernet	Ethernet port 10G	3.8W par Gbps
Niveau OTN	OTN port 10G	3.4W par Gbps
Niveau WDM	terminaux WDM, 40 canaux	230W par pair
	OLA, long span 80 km	110W par unité

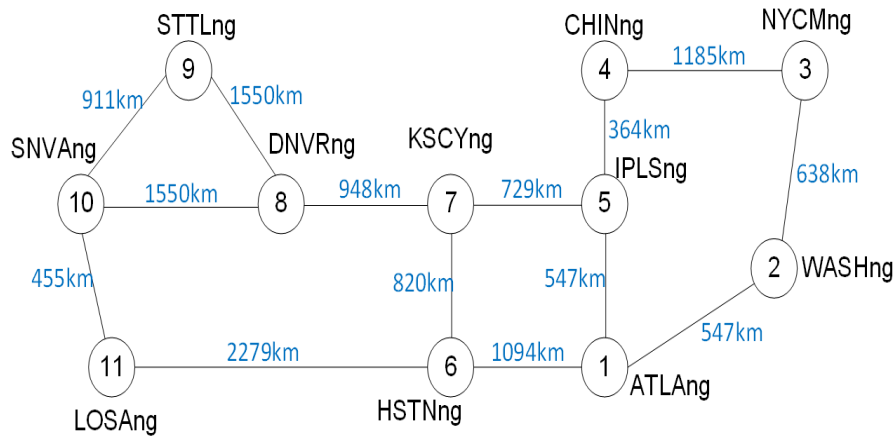


FIGURE 3.26 – Distance des liens dans Abilene

La consommation d'énergie de chaque amplificateur est de 110 watts. Nous supposons un multiplexage en longueur d'onde **WDM** de 40 canaux sur chaque liaison fibre et chaque liaison fibre optique est connectée avec une carte **IP Carrier Routing System 3 (CRS-3)** de 10Gbps.

La figure 3.27 illustre les résultats obtenus avec 1) la prise en compte de la consommation optique 2) la prise en compte de la consommation au niveau Ethernet 3) la prise en compte de la consommation Ethernet et optique.

Le résultat de la solution avec des coûts uniquement relatifs aux amplificateurs optique correspond à celui obtenu avec la solution traditionnelle (Legacy) et n'apparaissent pas différenciés sur la figure. Ceci s'explique par la topologie du réseau Abilene où la distance des liens entre deux noeuds n'est pas plus importante que la somme des distances des liens d'un chemin de traverse entre les deux

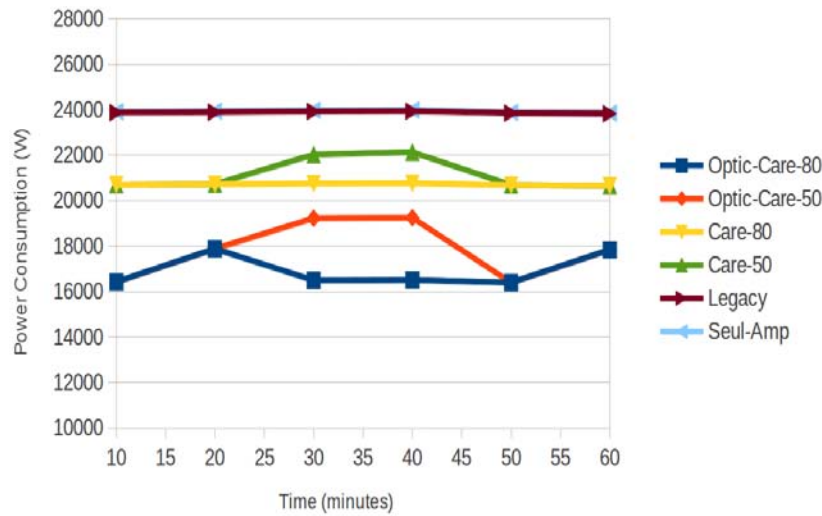


FIGURE 3.27 – Résultats d'évaluation

neouds. Le plus court chemin en nombre de noeud correspond au plus court chemin en distance. La solution combinée optique /Ethernet obtient de meilleurs résultats.

Remarque : Aux temps 20 minutes et 50 minutes, le trafic est supérieur à 50% ce qui explique les meilleurs performances du seuil $X = 80$ devant celui de $X = 50$.

Conclusion : La métrique obtenant les meilleurs résultats est la métrique Optic care 80 ($X = 80$).

Cas de réseaux hétérogènes

Nous avons défini une ingénierie verte spécifique à Ethernet et évaluée celle-ci sur des liens de débits similaires. Dans la mesure où les réseaux sont souvent hétérogènes, nous considérons cet aspect en indiquant comment la politique proposée peut être appliquée ou étendue à des réseaux hétérogènes. Des travaux complémentaires à ceux menés dans cette thèse sont nécessaires pour approfondir ce problème.

Considérons les cas d'hétérogénéité Ethernet ci-après.

Cas de liens Ethernet à débits différents :

Les modèles de coût pour Ethernet sont similaires sur des liens à 1Gbps et 10Gbps, la valeur des métriques 25% reste inchangée. Seule la valeur de charge sera différente. Ainsi les liens à 10 Gbps seront privilégiés.

Cas de liens Ethernet avec mécanisme d'endormissement et sans mécanismes d'endormissement :

Les liens sans économie d'énergie seront privilégiés. Dans la mesure où ils ne sont pas source d'économie d'énergie autant y faire passer le trafic. Leur coût est de 1. Les autres liens ont des coûts fonction de la métrique, [EAGER](#) ou [CARE](#).

Cas de liens à Economie d'énergie en technologie Ethernet et autre que Ethernet :

Dans ce cas nous proposons de définir pour chaque technologie la valeur de paramètre idoine, 25% pour Ethernet, 10% par exemple pour un lien optique.

3.5 Conclusion

En s'appuyant sur l'analyse de consommation de puissance effectuée sur Ethernet 802.3 az, nous avons proposé une ingénierie verte qui via le plan de contrôle étend l'économie d'énergie réalisée par le standard. Nous en avons évalué les performances par expérimentation et simulation.

La mise en place d'une ingénierie ne se justifie pas pour des réseaux très faiblement chargés dans la mesure où les liens se mettent fréquemment en sommeil. Elle est avantageuse pour des réseaux, moyennement, à fortement chargés avec en contrepartie une pénalité de délai sans toutefois occasionner plus de congestion qu'un routage standard.

L'originalité de notre démarche est de lier l'ingénierie au modèle de coût en énergie, et de prendre en compte l'occurrence de congestion. Plus précisément nous avons proposé 3 métriques de routage, qui permettent de faire des compromis entre délai et gain d'énergie avec des mises en oeuvre de complexité différentes.

La politique [CARE](#) 80 qui repose sur des changements gradués des poids des liens et évite de trop charger les liens en assignant un poids important dès lors que la charge dépasse 80% fournit un bon compromis entre le gain en économie d'énergie et le délai.

La même mécanique appliquée à un réseau optique en prenant en compte le coût des amplificateurs sur le lien fournit également de bons résultats.

L'inconvénient de cette méthode est lié à la notion de mesure de charge. Les poids sont fonction de la charge alors que dans la politique [EAGER](#), la charge est comparée à un seuil pour décider entre deux valeurs de poids. Il peut donc être préférable d'utiliser la politique [EAGER](#) plus simple à réaliser, d'autant plus que cette métrique fournit de bons résultats sans pénalité de délai pour des réseaux faiblement à moyennement chargés. Les notions de faible, moyenne, et fortement découlent du modèle de puissance du standard IEEE 802.z analysé dans le chapitre précédent. Les valeurs que nous avons utilisées peuvent être modifiées dans la définition des métriques pour s'adapter à différents équipements constructeurs.

Les travaux menés dans ce chapitre mettent en avant un problème de mesure de charge. La performance en économie d'énergie est liée à la charge du réseau. La notion de charge est quant à elle liée à une période de mesure. Ainsi le réseau Abilene qui a une charge moyenne de 2% et peut donc

être considéré comme très faiblement chargé (inférieur à 10%) présente néanmoins des périodes de charge fluctuant entre 10% et 20% avec quelques pointes au-delà de 50%, ce qui explique en ce cas l'intérêt d'une ingénierie verte. La mesure de charge que nous avons prise correspond à la périodicité des échanges d'annonces de routage. Une périodicité faible permet de suivre finement la charge mais génère un surplus de message de routage et peut conduire à des problèmes d'oscillations de route.

CHAPITRE 4

Ingénierie verte centralisée par réseau programmable SDN

Sommaire

4.1	Présentation de SDN (Software Defined Network)	100
4.1.1	L'architecture de SDN	101
4.1.2	Principaux Eléments	102
4.1.3	OpenFlow	104
4.2	SDN et ingénierie de trafic	107
4.2.1	Principe du Segment routing	108
4.2.2	Partage de charge et Ingénierie verte.	110
4.3	Evaluation	112
4.4	CONCLUSION	117

Ce chapitre traite de la mise en oeuvre de l'ingénierie verte, il réalise une preuve de concept en environnement centralisé en se posant la question d'améliorer les politiques proposées par un partage de charge.

Nous avons proposé et analysé les politiques d'ingénierie verte et concluons notre travail par l'étude de leurs mises en oeuvre. Actuellement, l'ingénierie de trafic est mise en oeuvre dans les réseaux opérateurs essentiellement à partir de technologie [MPLS](#) et à partir d'une administration centralisée ([Multiprotocol Label Switching - Transport Profile \(MPLS-TP\)](#)). Cependant l'avènement d'un plan de contrôle intelligent caractéristique des nouvelles architectures de réseau programmable, est un cadre de mise en oeuvre prometteur.

Dans ce chapitre nous examinons la faisabilité de nos travaux en réalisant une preuve de concept à base de [SDN](#). De plus nous montrons comment les résultats obtenus par les politiques peuvent être complétés par une politique d'ingénierie bien connue pour un objectif de qualité de service, le partage de charge.

4.1 Présentation de SDN (Software Defined Network)

Le réseau programmable (Software Defined Network) représente une évolution majeure dans le domaine de la conception et de la gestion de réseaux. Il concrétise des travaux précurseurs ayant eu pour objectif d'augmenter l'adaptation des réseaux en les rendant programmables, avec deux axes de travaux : ceux sur la programmation, les réseaux actifs et ceux sur l'architecture séparant le contrôle et la donnée, [MPLS](#).

L'approche par réseaux actifs développée dans les années 90 pour intégrer de nouveaux services dans le réseau internet est à l'origine du réseau programmable [[Che00](#)] [[PCB01](#)]. Par cette approche les traitement des flux peuvent être programmés sur le réseau, grâce à une interface de programmation, ou encore un langage tel [Packet Language for Active Networks \(PLAN\)](#) permettant d'envoyer et l'information et le traitement à effectuer sur cette dernière par les routeurs [[HMA⁺99](#)]. Bien que cette recherche ne se soit pas concrétisée (raison de coûts, absence de standards, prévalence des aspects sécurité), elle est à l'origine de nombreuses avancées dans les domaines de la virtualisation de réseaux, l'utilisation de fonction programmables, la vision d'architecture unifiée.

Les travaux sur la séparation du plan de contrôle et des données ont été initiés dans les années 90 pour faire face à l'augmentation des volumes de transfert de données en accélérant les traitements de routeurs. Les fonctions du plan de contrôle, l'intelligence du réseau, étaient effectuées par logiciel

alors que le plan de donnée, le relayage, effectué par le matériel était nettement plus rapide. Ces travaux à l'origine de [MPLS](#) se sont poursuivis avec par exemple en 2004, [Routing Control Platform \(RCP\)](#) [CCF⁺05], qui collectait de façon centralisée des données sur la topologie du réseau pour choisir des routes à la place des routeurs puis leur communiquer, via le protocole iBGP ([Internal Border Gateway Protocol \(IBGP\)](#)), afin qu'ils effectuent leur relayage IP.

L'approche [SDN](#) intègre ces deux approches, elle sépare le plan de données du plan de contrôle en offrant la possibilité de programmer les fonctions de contrôle. Elle s'appuie sur la dissociation du système qui prend une décision sur l'endroit où le trafic doit être envoyé et sur le système qui envoie les informations, avec une communication standard entre les deux systèmes, et des interfaces permettant de programmer le système de décision pour obtenir un réseau adaptable à des services.

4.1.1 L architecture de SDN

L'[Open Network Foundation \(ONF\)](#) est la principale autorité de standardisation des réseaux [SDN](#). En proposant des standards ouverts, elle simplifie la conception et la maintenance du réseau dans la mesure où les instructions sont fournies par le contrôleur [SDN](#) et non pas par des protocoles et dispositifs spécifiques à chaque vendeur. L'architecture de réseau [SDN](#) proposée par l'[ONF](#) est illustrée en figure 4.1

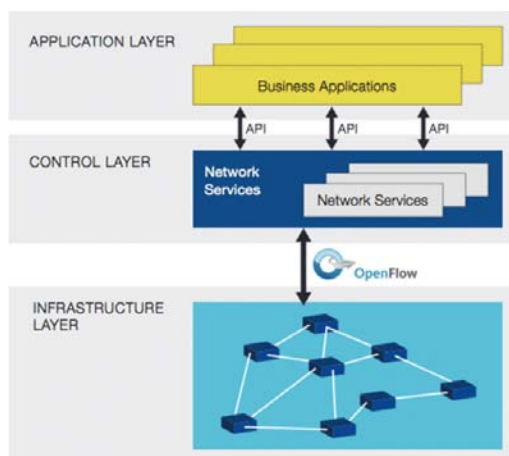


FIGURE 4.1 – Architecture de réseau SDN [ONF]

Nous retrouvons un plan application, qui grâce aux interfaces de programmation, spécifie dans le plan de contrôle, les services réseaux qui seront exécutés par les équipements formant l'infrastructure du réseau. Ceux-ci sont des équipements de relayage, nommés switch, ils relayent les informations en examinant leur entête.

4.1.2 Principaux Eléments

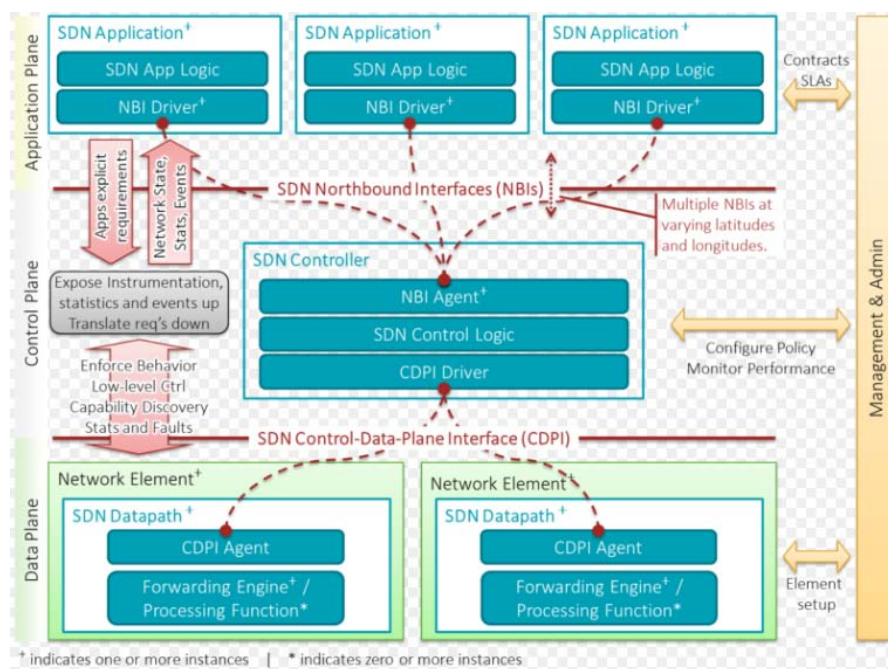


FIGURE 4.2 – Principaux Eléments SDN [M⁺13]

L'IEEE s'appuie sur l'ONF pour spécifier le SDN en précisant le plan contrôle par la figure 4.2. Les éléments de l'architecture sont :

- les interfaces
- Le contrôleur
- les switch

Les interfaces (ou **Application Programming Interface (API)**)

Elles sont au nombre de deux. L'interface Nord (Northbound [API](#)) est utilisée pour la communication entre les applications et les services. Le comportement d'un switch est contrôlé à distance grâce à l'interface Sud (Southbound [API](#)) qui est l'interface entre le plan de contrôle et le plan de données sur l'élément contrôleur. La plus populaire des interfaces est OpenFlow ; (des chercheurs ont proposé DevoFlow comme autre solution d'interaction switch-contrôleur.)[[CKY11](#)]

Le contrôleur

Le principe de base des réseaux [SDN](#) c'est d'utiliser un contrôleur logiquement centralisé pour contrôler à distance tous les switchs du réseau. Plusieurs propositions de contrôleur physiquement centralisé tel que POX, RIU, NOX et Maestro [[CCN11](#)] ont été faites. Il leur est reproché leur vulnérabilité face à la défaillance, et des problèmes de passage à l'échelle. Pour résoudre ces points, la communauté [SDN](#) a proposé des contrôleurs logiquement centralisés mais physiquement distribués. Il peut y avoir plusieurs contrôleurs où chacun gère une partie du réseau mais communiquent entre eux pour garder une vue globale et donc un contrôle centralisé. ElastiCon, HyperFlow et Onix [[KCG⁺10](#)] sont autant d'exemples de contrôleurs distribués.

Les switch

Les switch sont des équipements de relaying qui examinent l'entête d'informations entrantes puis en fonction de la logique du plan de contrôle exécutent des actions. L'entête est conforme à celle produite dans un réseau Internet, avec une encapsulation caractéristique de chaque protocole. La définition de l'entête qui est traitée par le switch est variable, ce peut être une entête de niveau 2 de type Ethernet aussi bien qu'une entête de niveau 4 Transport. Les actions sont mémorisées dans des Tables.

Dans cette thèse nous utilisons le switch standard, le switch Openflow. Il est constitué d'une ou plusieurs tables de flux, construites à partir de mémoire adressable par contenu ternaire ([TCAM](#)), d'une entité de relaying de paquets et d'un canal sécurisé (qui est en fait une interface) vers le contrôleur externe. Le canal permet d'échanger des commandes et des paquets (Packet_In : du switch au contrôleur, et Packet_Out : du contrôleur au switch) entre le switch et le contrôleur. Le contrôleur configure et gère le switch à travers le canal sécurisé.

4.1.3 OpenFlow

Le protocole OpenFlow gère l'échange de messages de contrôle entre le contrôleur et switch. Les messages openflow permettent d'installer les règles de fonctionnement aux switches (exemple : messages FlowMod) et d'obtenir des informations de topologie et de trafic.

Explicitons les messages avant d'en illustrer l'utilisation et l'aspect performance.

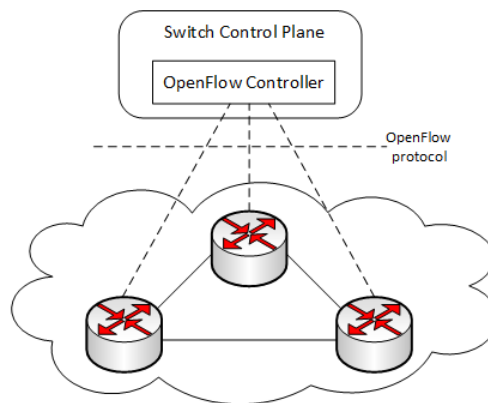


FIGURE 4.3 – Architecture OpenFlow

Il existe 3 types de messages dans le protocole OpenFlow : les messages 1) contrôleur- switch, 2) asynchrones et 3) symétriques.

Les messages contrôleur-switch :

ils sont initiés par le contrôleur et servent à gérer directement ou à vérifier l'état du switch. Ces messages ne nécessitent pas forcément de réponse de la part du switch.

- Features_Request or Reply : le contrôleur peut demander les caractéristiques du switch en envoyant un message Features_Request. Le switch doit répondre avec un message Features_Reply qui liste ses caractéristiques. Cet échange est communément effectué lors de l'établissement du canal OpenFlow.
- Configuration : ces messages permettent au contrôleur de définir et de demander les paramètres de configuration du switch.

- **Modify_State** : ces messages permettent au contrôleur de gérer l'état du switch. Leurs objectifs principaux sont d'ajouter, de modifier ou de supprimer les entrées de flux de la table du switch, ainsi que de gérer les propriétés des ports du switch.
- **Read_State** : ces messages permettent au contrôleur de collecter les statistiques de trafic des tables de flux et de ports, présentes dans le champ Counter.
- **Barrier_Request or Reply** : ces messages permettent au contrôleur de recevoir des notifications lorsque des opérations ont été achevées.
- **Packet_Out** : ces messages sont utilisés par le contrôleur pour envoyer des paquets sur des ports précis, et également pour forwarder les paquets qui ont engendré un envoi de **Packet_In** par le switch.

Les messages asynchrones :

sont initiés par le switch pour informer le contrôleur de la survenance d'événements (arrivée d'un paquet, erreur, ...) et de changements d'états du switch.

- **Flow_Removed** : ces messages indiquent au contrôleur qu'une entrée de flux a été supprimée dans la table de flux du switch. Ils précisent si la suppression a été due à l'expiration du délai d'inactivité (idle timeout), à l'expiration du délai d'attente maximale (hard timeout), ou à la réception d'un message **Modify_State**, envoyé du contrôleur au switch.
- **Port_Status** : ces messages indiquent une modification dans la configuration des ports du switch.
- **Error** : ces messages permettent au switch de notifier le contrôleur d'une erreur quelconque.
- **Packet_In** : un message est envoyé au contrôleur pour chaque paquet qui ne correspond à aucune entrée de flux, ou lorsque l'action **CONTROLLER** est associée à l'entrée de flux à laquelle est associée le paquet. Si le switch possède assez de mémoire interne pour mémoriser le paquet qui a entraîné l'envoi d'un **Packet_In**, ce-dernier ne contient que l'entête du paquet et un identifiant de buffer. Cependant, si le switch ne possède pas de buffer interne, ou s'il n'a plus de mémoire interne de libre, le message **Packet_In** contient le paquet entier.

Les messages symétriques :

Les messages symétriques sont envoyés sans sollicitation, par n'importe lequel des deux équipements (soit le contrôleur, soit le switch).

- Hello : ces messages sont échangés entre le switch et le contrôleur lors de la négociation de la connexion.
- Echo_Request or Reply : ces messages sont envoyés par l'un ou l'autre des équipements, et attendent en réponse un message Echo_Reply. Ils peuvent servir à calculer la latence et la bande passante de la connexion et à vérifier que la connexion switch contrôleur est toujours opérationnelle.

Une illustration de l'utilisation des messages OpenFlow est fournie en Figure 4.4.

Le contrôleur configure les règles du plan de données par le message Flowmod, en écrivant dans les tables FlowTable qui ont une entrée par flux, comment le switch doit traiter le flux :

- 1) Lorsqu'une donnée est reçue par le switch qui ne correspond à aucune entrée de la table, 2) un message de Packet-in incluant l'entête de la donnée est transmis au contrôleur. 3) Celui-ci répond par un message Flowmod, 4) l'action est mémorisée dans la table 5) le paquet est relayé. Le processus est répété au switch suivant : 6), 7), 8), 9).

La découverte de la topologie est établie en environnement Ethernet par un protocole [LLDP](#), ou bien un protocole d'arbre recouvrant ([Spanning Tree Protocol \(STP\)](#)), qui permet au contrôleur de remplir une table de la topologie.

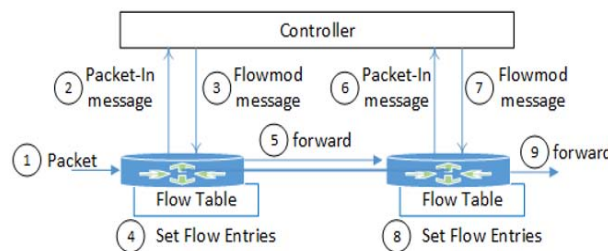


FIGURE 4.4 – Échange de messages Openflow

Conclusion : L'installation des flux que nous venons d'illustrer est un paramètre majeur de la

performance du SDN que ce soit en raison de limitations de la taille des tables qui devraient être levées par des avancées technologiques, ou en raison des échanges protocolaires. Ceux-ci dépendent du nombre de flux à identifier, de la fréquence d'adaptation de l'ingénierie et du nombre de switch. Nos travaux prennent en compte ce dernier point.

4.2 SDN et ingénierie de trafic

Le SDN est conçu pour faire de l'ingénierie de trafic simplement. La connaissance globale du réseau en un point central facilite la prise de décision. Les nombreux projets menés par la communauté de recherche et les opérateurs attestent de ses possibilités cependant il subsiste des limitations de passage à l'échelle et de consommation de bande passante :

- La virtualisation entraîne l'ajout d'un entête conséquente dans les messages.
- La virtualisation entraîne également l'apparition de latence lorsque le switch est opéré en mode réactif. Dans ce mode les règles de fonctionnement sont installées à l'apparition d'un nouveau flux par opposition au mode statique, où les tables sont préinstallées.
- L'incorporation de l'ingénierie de trafic au SDN requiert pour chaque switch de communiquer avec le contrôleur central pour chaque flux de trafic.

Afin de diminuer les échanges en SDN, nous proposons :

- 1) d'utiliser un routage par la source, où le routeur source obtiendrait l'entête indiquant le chemin à suivre du contrôleur et où les actions à mener sur l'entête seraient mémorisées sur les switch. Le nombre d'échanges entre switch et contrôleur n'est alors plus dépendant du nombre de switch du chemin.
- 2) une ingénierie via MPLS. L'intérêt de ce mécanisme par identifiant de flux est de diminuer le nombre d'entrée de table par rapport à une identification IP sur adresse destination.

L'inconvénient du mécanisme de chemin contraint classique (à partir des protocoles Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE) ou Constraint-based Routing Label Distribution Protocol (CR-LDP) pour établir les chemins MPLS, les Label Switch Path) est son manque de souplesse, il faut indiquer explicitement les éléments de la route à suivre, qui peut être longue, et

la non autonomie des équipements, l'équipement n'a qu'une action possible qui est de relayer vers l'élément identifié dans le champ idoine.

C'est pourquoi nous proposons de s'appuyer sur le protocole Segment Routing en cours de standardisation par l'IETF) qui permet de remédier à ces problèmes en considérant le chemin non pas comme une suite d'identifiants d'équipements mais comme une suite de labels où chaque label correspond à une action.

Nous étudions dans cette section la mise en oeuvre de l'énergie verte définie dans le chapitre précédent, par cette méthode. Présentons en premier lieu, le Segment Routing avant de détailler la mise en oeuvre.

4.2.1 Principe du Segment routing

Le Segment Routing est un routage par la source qui identifie une route par un ensemble de segments connus par leurs identifiants, SID ([Segment Identifiers \(SID\)](#)). La route est calculée par un contrôleur de route qui la transmet aux éléments du réseau. Le nœud d'entrée rajoute aux informations à transmettre une entête constituée d'un ensemble d'identifiants de segments par lesquels l'information va circuler.

Un identifiant de segment entraîne une action de type relaiage ou de type service. Avec ce routage, il n'y a plus besoin d'utiliser de protocoles d'établissement de chemin distribué, les chemins et leurs labels sont prédéfinis.

Modes d'utilisations du segment routing

Le mode du routage du segment routing est soit un mode de tunnel, où le trafic est émis sur un chemin défini par ses segments, ou bien un mode autonome avec un routage au plus court chemin et répartition de charge.

Opérations sur les segments

- Les opérations effectuées par un nœud de relaiage sont :
 - **PUSH** : Insertion d'un segment à la tête de la liste des segments.

- **NEXT** : Le segment actif est terminé ; le segment suivant devient actif.
- **CONTINUE** : Le segment actif n'est pas terminé et donc reste actif.

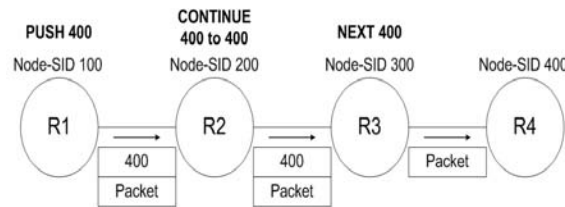


FIGURE 4.5 – Opérations du segment routing

Un exemple simple de fonctionnement de ces opérations est présenté Figure 4.5. Le routeur R1 transmet un paquet à destination du routeur R4 en rajoutant au paquet l'identifiant de segment du routeur R4 (400, le chemin pour joindre R4 dont l'identifiant est 400). Le routeur R2 exécute l'opération CONTINUE puis l'avant dernier routeur enlève le SID du segment qui se termine.

Le concept de segment permet de diriger finement le trafic, grâce à un empilement d'identifiants similaires à celui de MPLS pour les tunnels de trafic (Figure 4.6), tout en effectuant du partage de charge, [Equal Cost Multi Path \(ECMP\)](#) [FM15] [FPD⁺16] entre la source et la destination comme en Figure 4.7

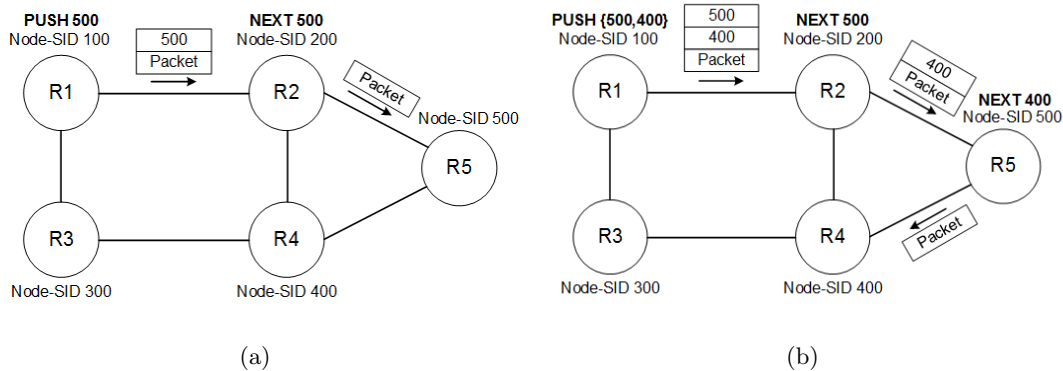


FIGURE 4.6 – Empilement de labels en segment routing

- Configuration via [SID \(Tube\)](#) : Le contrôleur assigne le SID du noeud d'entrée du tunnel.
- Configuration via Shortest Path : Le contrôleur ne définit pas d'entrée du tunnel puis un

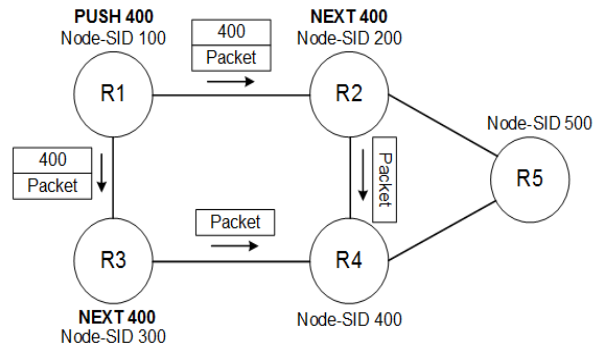


FIGURE 4.7 – Illustration de partage de charge dans ECMP

flux est dirigé par de multiples chemins de coût identiques ([ECMP](#) [[FM15](#)] [[FPD⁺16](#)]) entre la source et la destination (Fig 4.7).

L'intérêt de la répartition de charge est d'améliorer l'utilisation des ressources des opérateurs tout en évitant les congestions, un objectif de qualité de service. Cet objectif nous semble particulièrement pertinent dans la mesure où il réduit les délais mais cependant il peut augmenter la consommation d'énergie, dans la mesure où le trafic transite par un nombre plus importants de liens. Néanmoins les lignes sont alors moins chargées et peuvent donc se mettre en sommeil.

Nous nous proposons d'évaluer l'intérêt de ce mécanisme.

4.2.2 Partage de charge et Ingénierie verte.

Pour mettre en place l'ingénierie verte, nous exprimons deux méthodes, une avec et l'autre sans partage de charge. De façon à ne pas trop pénaliser, le partage de charge, sur l'énergie dépensée, nous proposons de ne pas considérer celui-ci que sur une partie du réseau, la topologie verte.

Nous avons montré dans le chapitre précédent l'intérêt du routage vert devant le routage au plus court chemin, la question que nous nous posons à présent est de savoir si l'utilisation du partage de charge ([ECMP](#)) peut améliorer les politiques de direction de trafic [CARE](#) et [EAGER](#) qui elles agrègent les trafics.

Précisons en premier lieu, la mise en oeuvre des politiques en [SDN](#).

4.2.2.1 Mises en oeuvre des politiques par métriques et partage de charge. Les deux politiques que nous proposons d'examiner sont définies par les processus ci-après :

- 1) Calcul du routage vert sur la topologie complète et direction du trafic, par tunnel sur des routes prédéfinies comme dans le chapitre précédent.
- 2) Calcul du routage vert pour en déduire une topologie verte grâce à la vision centralisée du [SDN](#) puis application du routage classique au plus court chemin et partage de charge.

Nous ne considérons pas le partage de charge sur tout le réseau mais uniquement sur les liens verts ce qui permet de mettre des liens en sommeil, puis de répartir le trafic.

Le fonctionnement est cyclique. Dans les évaluations nous prenons un intervalle de 5 minutes (correspondant aux mesures de la matrice Abilene).

Calcul du routage vert

- a) Le contrôleur collecte cycliquement les données de trafic via les messages openflow statistiques, ainsi que la topologie par sa procédure de découverte. Nous utilisons [LLDP](#) (802.3ab).
- b) Il en déduit la charge des liens et calcule le routage vert pour chaque couple entrée destination du réseau.
- c) Les chemins calculés sont associés à des identifiants de segment qui sont installés dans la table de flux des switch.
- d) Les flux de trafic sont transmis dans ces tunnels.
- Les liens ne supportant pas de trafic se mettent en sommeil conformément au mécanisme [EEE](#) 802.3az sans intervention du contrôleur.

Calcul du routage [ECMP](#) sur topologie verte

- a) Le contrôleur effectue également le calcul de routage vert.

- b) Il en déduit une topologie verte qui exclut les liens qui ne sont pas chargés, et obtient une topologie virtuelle constituée de liens éveillés.
- c) les chemins sont calculés par [ECMP](#) sur la topologie virtuelle.
- d) les flux de trafic sont émis avec un label.

Plutôt que d'agréger les flux par tunnel comme c'est le cas avec les politiques vertes, le routage par répartition de charge est appliqué sur la topologie virtuelle.

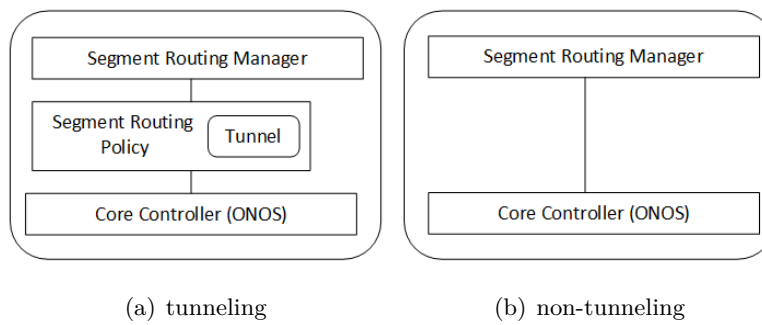


FIGURE 4.8 – Architecture de l'ingénierie verte en SDN et segment routing

4.3 Evaluation

Nous évaluons l'intérêt de la méthode par tunnel d'ingénierie verte devant la méthode par répartition de charge par émulation de [SDN](#) sous Mininet (le simulateur [NS3](#) ne dispose pas de module [SDN](#)). Nous avons émulé la topologie Abilene avec des liens IEEE 802.az à 10Gbps, présentée dans le chapitre précédent. Pour des raisons de temps d'exécution, nous avons réduit la matrice de trafic en ne considérant que les cinq dernières minutes de l'heure [[Yin](#)] pendant une période de 24 heures.

Le Contrôleur est un contrôleur [Open Network Operating System \(ONOS\)](#) avec lequel nous avons un module segment routing.

Les algorithmes que nous avons implantés sont indiqués ci-contre, (algorithme 2, algorithme 3)

Les métriques de comparaison sont celles utilisées précédemment, à savoir : l'économie d'énergie, l'utilisation des liens et le délai.

Algorithm 2 : Tunneling TE

```

1  $X = (1, 1, \dots, 1) \rightarrow$  Initially all the link are active
2  $F = (f_1, f_2, \dots, f_N) \rightarrow$  Total flows in Network =  $N$ 
   /*Phase1-Green Metrics and pre-established paths*/
3 for each flow  $f_i \in F (i = 1, 2, \dots, N)$  do
4    $S(i) \leftarrow$  Pre-established Path (calculate with EAGER or CARE metrics)
5    $P_{xy}(i) \leftarrow$  Select Path  $(x, y)$  (denote Path between source  $(x)$  and destination  $(y)$ )
6    $P \leftarrow P_{xy}(i)$ 
7 end
   /*Phase2-Traffic Engineer with Tunneling*/
8 for each flow  $f_i \in F (i = 1, 2, \dots, N)$  do
9   Tunnels are set follow Pre-established Path  $P_{xy}(i)$ 
10  Controller tells every Segment Routers
11  Segment Router updates and uses tunnel table to transfer traffic  $f_i$ 
12 end

```

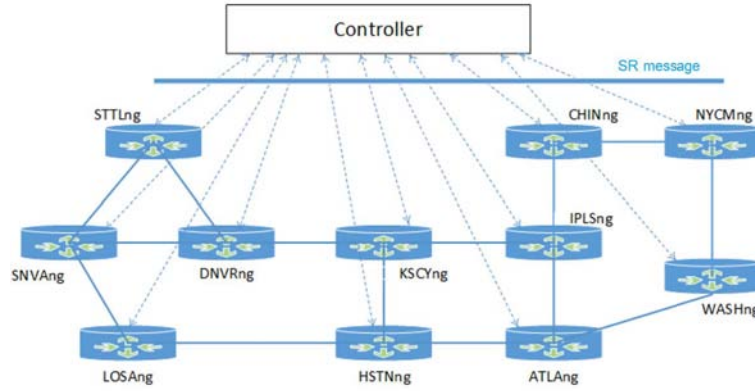


FIGURE 4.9 – Topologie expérimentale SDN-Abilene

Dans le chapitre précédent nous avons évalué l'économie d'énergie en simulant le fonctionnement du lien IEEE 802.3 az, dans ce chapitre nous utilisons l'outil d'émulation qui ne dispose pas du module Ethernet. Afin de réduire les développements nous n'avons pas créé de module Ethernet.

Algorithm 3 : Non-tunneling TE

```

1  $X = (1, 1, \dots, 1) \rightarrow$  Initially all the link are active
2  $F = (f_1, f_2, \dots, f_N) \rightarrow$  Total flows in Network =  $N$ 
   /*Phase1-Green Metrics and Selecting links to sleep*/
3 for each flow  $f_i \in F (i = 1, 2, \dots, N)$  do
4    $S(i) \leftarrow$  Pre-established Path (calculate with EAGER or CARE metrics)
5    $P_{xy}(i) \leftarrow$  Select Path  $(x, y)$  (denote Path between source  $(x)$  and destination  $(y)$ )
6    $P \leftarrow P_{xy}(i)$ 
7   for each link  $l_k \in L (k = 1, 2, \dots, N)$  do
8      $X_k \leftarrow (P_{xy}(i), l_k)$ 
       Sleep links are selected by Kronecker delta function ( ).
       Kronecker delta function returns 1 if the link  $k$  belong to  $P_{xy}(i)$  and 0 otherwise.
       keep active link = 1 and sleep link = 0 at  $X_k \in X (k = 1, 2, \dots, L)$ 
9   end
10  Update( $X$ )
11 end
   /*Phase2-Traffic Engineer with Non-Tunneling*/
12 Green Virtual Topology is set with sleep links
13 Controller uses green virtual topology and controller tells every SR routers
14 Every router changes the forwarding table to route over the new paths

```

Nous avons évalué l'économie d'énergie au contrôleur en se référant aux travaux menés et en utilisant une table de correspondance charge/consommation d'énergie. Le contrôleur obtient les statistiques d'utilisation des switches avec la charge des liens, en déduit la consommation qui est ensuite traitée pour en déduire une économie d'énergie.

Résultats

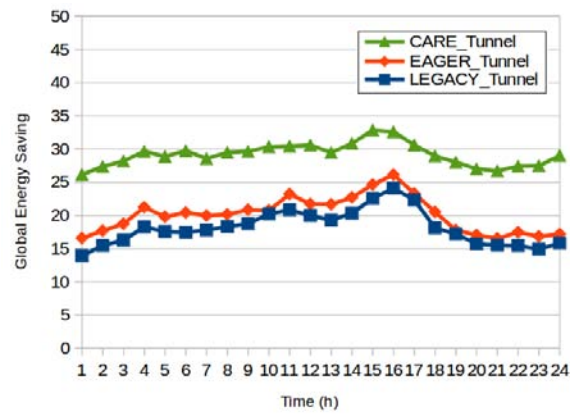
Les politiques sont comparées en les positionnant avec la référence de la politique standard qui émet au plus court chemin sans se soucier de la consommation (LEGACY).

Nous retrouvons comme précédemment de meilleurs résultats avec la méthode CARE 80 qu'avec

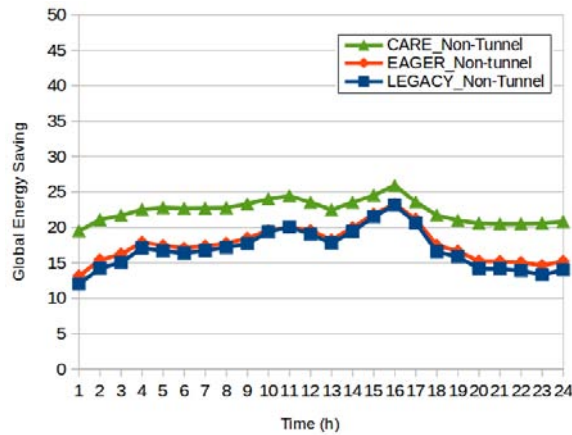
les autres méthodes et nous constatons que la politique de répartition de charge n'apporte rien en économie d'énergie. Le gain varie de 25% à 32% en dirigeant le trafic sur un chemin alors qu'il diminue avec la répartition (économie variant de 20% à 25%).

Concernant la congestion, le maximum d'utilisation de lien atteint 30% en CARE (contre 17% en legacy) en méthode tunnel et il est naturellement moindre en utilisant la répartition de charge (19%)

Quant au délai, il est pratiquement doublé par le tunnel, devant la méthode de base alors que le délai en répartition de charge est quasi identique.

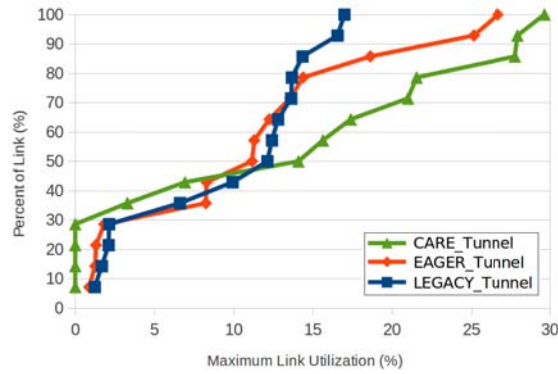


(a) tunnel

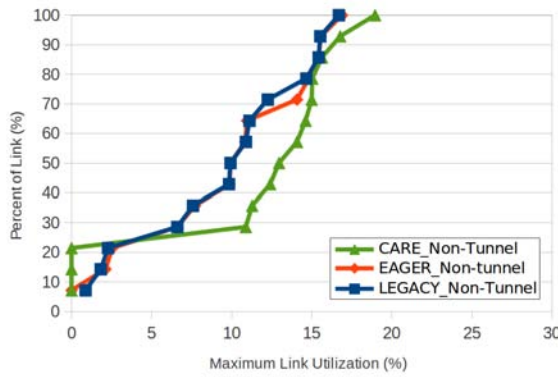


(b) partage de charge

FIGURE 4.10 – Economie d'énergie sur réseau SDN-Abilene en segment routing (a) tunnel, (b) partage de charge



(a) tunnel



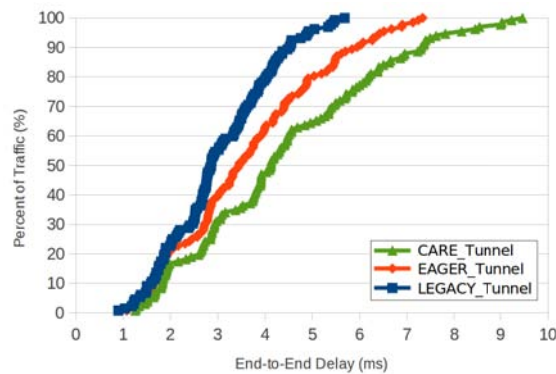
(b) partage de charge

FIGURE 4.11 – CDF de l'utilisation maximale de liaison sur réseau SDN-Abilene en segment routing
(a) tunnel, (b) partage de charge

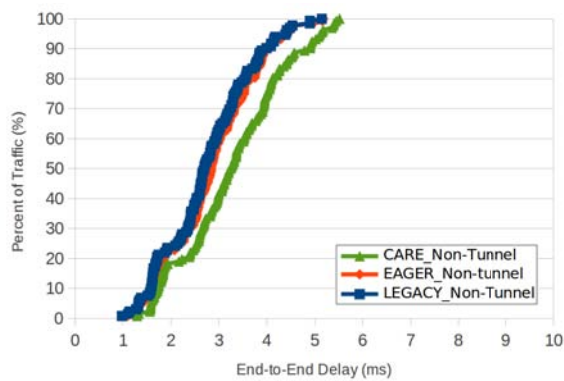
Conclusion

La méthode CARE complétée par le partage de charge permet d'économiser de l'énergie sans pénalité de délai.

Le gain obtenu par notre proposition est quasi doublé par rapport au standard IEEE 802.3az.



(a) tunnel



(b) partage de charge

FIGURE 4.12 – CDF de délai de bout en bout sur réseau SDN-Abilene en segment routing (a) tunnel, (b) partage de charge

4.4 CONCLUSION

Nous avons effectué une preuve de concept de nos travaux, en les mettant en oeuvre dans une architecture de contrôle centralisée, en technologie [MPLS](#) très répandue chez les opérateurs.

La réalisation centralisée par réseaux programmable [SDN](#) pose un problème de passage à l'échelle du nombre d'équipements que nous proposons de réduire par protocole de routage par la source intelligent, le segment routing, qui est en cours de standardisation par l'[IETF](#).

Le segment routing intègre une politique de partage de charge qui a l'avantage lorsqu'elle est combinée avec l'utilisation de notre proposition d'ingénierie de pouvoir économiser de l'énergie tout

en préservant la qualité de service, comme en atteste l'évaluation que nous avons présentée.

Nous avons effectué des travaux d'évaluation qui pourraient être étendus et formalisés analytiquement afin de mieux appréhender l'intérêt des politiques.

Ce chapitre qui conclut nos travaux mettent également en perspective des besoins d'analyse plus poussée concernant la surveillance de réseaux en [SDN](#), thèmes d'étude appelé à se développer comme en attestent les travaux en phase de démarrage sur le réseau de la recherche et de l'éducation Géant 4 dans le cadre du programme européen H2020.

Conclusion et Perspectives

Les travaux menés durant cette thèse ont porté sur l'étude des mécanismes permettant d'économiser l'énergie en réseau filaire en se préoccupant plus particulièrement du réseau Ethernet, pour lequel nous avons disposé d'un environnement standardisé avec la possibilité d'effectuer des expérimentations. Nous avons choisi d'économiser de l'énergie en mettant en sommeil des liens de réseau grâce au standard Ethernet IEEE 802.3az que nous avons modélisé et évalué ce qui nous a permis de mettre en avant l'importance de la charge de trafic à transmettre dans la quantité d'énergie consommée avec une croissance de consommation très rapide en fonction de la charge.

En nous appuyant sur ce résultat nous avons proposé d'étendre le gain obtenu par le standard en développant une politique d'ingénierie de trafic que nous avons qualifiée de verte, pour indiquer son objectif d'économie d'énergie. Cet objectif est atteint en envoyant sur les liens suffisamment peu de trafic pour leur permettre de s'endormir. Notre approche est une approche générique qui peut s'adapter à plusieurs systèmes dès lors qu'ils utilisent un mécanisme de mise en sommeil. La détermination de la quantité de trafic à envoyer est fonction du modèle de puissance et de la technologie dont nous déduisons un modèle de coût.

Dans le cadre d'Ethernet nous avons étudié deux modèles de coûts pour refléter son modèle de puissance, [EAGER](#) et [CARE](#). Le premier répartit le trafic en tentant de laisser les liens quasiment vides, alors que le second répartit davantage le trafic en considérant qu'un lien peut encore s'endormir et donc fournir un gain d'énergie si sa charge n'est pas trop importante. Le choix d'un modèle devant un autre est lié à ses performances et à sa complexité de mise en oeuvre.

En ce qui concerne la performance nous avons considéré l'économie d'énergie mais également les pertes pouvant résulter d'une surcharge trop importante, ainsi que le délai qui peut être accru de

par l'emprunt d'un chemin de réseau peu chargé mais long. Nous avons mené nos évaluations par simulation et effectué des expérimentations. Les deux modèles obtiennent des résultats intéressants avec cependant un avantage au deuxième modèle en performance et en qualité de service, car nous avons également incorporé à ce modèle un mécanisme permettant de ne pas surcharger les liens. En ce qui concerne la complexité du modèle liée à la détermination de la charge du lien, c'est-à-dire au mécanisme de mesure à mettre en place sur chaque élément, le modèle **EAGER** est avantagé devant le modèle **CARE** qui lui repose sur une valeur graduée de la charge et non une comparaison à un seuil.

Après avoir montré la pertinence de notre proposition en regard des gains obtenus par une étude de simulation et expérimentation, nous avons effectué une preuve de concept en considérant l'émulation d'une architecture de réseau logicielle (Software Defined Network). Si notre approche est utilisable dans un réseau Internet avec un routage **OSPF** il est cependant besoin d'en régler les paramètres (seuil de charge, seuil de congestion) sur chaque équipement ce qui peut s'avérer fastidieux. L'approche **SDN** qui centralise en un point et configure par logiciel l'ingénierie de trafic est une mise en oeuvre tout à fait adaptée à notre travail, comme nous l'avons montré dans notre mise en oeuvre. Nous avons en partie abordé les limitations de l'approche **SDN** en regard du passage à l'échelle, en développant l'ingénierie par un protocole de routage par la source de façon à diminuer le nombre d'échanges entre l'élément de décision et les équipements de relayage.

Perspectives

Le travail que nous avons mené est une première approche d'ingénierie de trafic économe, dont nous avons montré la faisabilité et l'intérêt. Cette approche devrait pouvoir être améliorée par des propositions de métriques différentes qui prendraient en compte la spécificité du trafic à transférer, ses caractéristiques de qualité de service, ou encore l'état global du réseau, avec par exemple une absence d'ingénierie verte en cas de réseau faiblement chargé ou au contraire très chargé.

L'étude de réseaux hétérogènes avec des mécanismes de sommeil à coûts différents ou encore d'autres mécanismes d'économie d'énergie est également une perspective de travail qui nous semble intéressante de poursuivre.

Une perspective plus large de notre travail concerne l'administration des réseaux logiciels [SDN](#), avec un besoin de surveillance et mesure. Les travaux que nous avons menés mettent en avant le besoin d'un mécanisme de mesure de charge, qui permette de mettre en place des politiques d'ingénierie verte, ou plus généralement les politiques d'ingénierie de trafic en architecture [SDN](#). Il nous semble important d'étudier plus avant l'impact du mécanisme de mesures de charge et de sa périodicité sur notre politique ou plus globalement d'étendre les possibilités du [SDN](#) pour offrir un ensemble d'outils adaptés.

Liste des publications

Articles de conférences internationales

1. C. Thaenchaikun, G. Jakllari, B. Paillassa and W. Panichpattanakul, Mitigate the Load Sharing of Segment Routing for SDN green Traffic Engineering , in *The 2016 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2016)*, IEEE, Phuket, Thailand, October 2016.
2. C. Thaenchaikun, G. Jakllari, and B. Paillassa, Augmenting the energy saving impact of IEEE 802.3az via the control plane , in *2015 IEEE International Conference on Communication Workshop (ICCW 2015)*, London, UK, June 2015, pp. 2843-2849.
3. B. Paillassa, R. Dhaou, G. Jakllari, and C. Thaenchaikun, Performance evaluation of energy efficient policies for ethernet switches , in *Wireless Communications and Mobile Computing Conference (IWCMC 2013)*, Cagliari, Sardinia, Italy, July 2013, pp. 797-802.

Bibliographie

- [ACG13] E. Amaldi, A. Capone, and L. G. Gianoli. Energy-aware ip traffic engineering with shortest path routing. *Comput. Netw.*, 57(6) :1503–1517, April 2013.
- [ADS09a] Asymmetric digital subscriber line 2 transceiver (ADSL2)-Extended bandwidth ADSL2 (ADSL2plus). *ITU-T G.992.5*, Jan 2009.
- [ADS09b] Asymmetric Digital Subscriber Line Transceivers 2 (ADSL2). *ITU-T G.992.3*, Apr 2009.
- [AFLV08] Mohammad Al-Fares, Alexander Loukissas, and Amin Vahdat. A scalable, commodity data center network architecture. *SIGCOMM Comput. Commun. Rev.*, 38(4) :63–74, August 2008.
- [AGE⁺04] S. Avallone, S. Guadagno, D. Emma, A. Pescape, and G. Ventre. D-itg distributed internet traffic generator. In *Quantitative Evaluation of Systems, 2004. QEST 2004. Proceedings. First International Conference on the*, pages 316–317, Sept 2004.
- [AK08] Ganesh Ananthanarayanan and Randy H. Katz. Greening the switch. In *Proceedings of the 2008 Conference on Power Aware Computing and Systems*, HotPower’08, pages 7–7, Berkeley, CA, USA, 2008. USENIX Association.
- [ALW⁺14] Ian F. Akyildiz, Ahyoung Lee, Pu Wang, Min Luo, and Wu Chou. A roadmap for traffic engineering in sdn-openflow networks. *Computer Networks*, 71 :1 – 30, 2014.

- [ARSG06] H. Anand, C. Reardon, R. Subramaniyan, and A. D. George. Ethernet adaptive link rate (alr) : Analysis of a mac handshake protocol. In *Local Computer Networks, Proceedings 2006 31st IEEE Conference on*, pages 533–534, Nov 2006.
- [BA99] Albert-Laszlo Barabasi and Reka Albert. Emergence of scaling in random networks. *Science*, 286(5439) :509–512, 1999.
- [BCRR12] A. P. Bianzino, C. Chaudet, D. Rossi, and J. L. Rougier. A survey of green networking research. *IEEE Communications Surveys Tutorials*, 14(1) :3–20, First 2012.
- [BDB⁺11] R. Bolla, F. Davoli, R. Bruschi, K. Christensen, F. Cucchietti, and S. Singh. The potential impact of green technologies in next-generation wireline networks : Is there room for energy saving optimization? *IEEE Communications Magazine*, 49(8) :80–86, August 2011.
- [BDP10] A. Botta, A. Dainotti, and A. Pescapé. Do you trust your software-based traffic generator? *IEEE Communications Magazine*, 48(9) :158–165, Sept 2010.
- [BDP12] Alessio Botta, Alberto Dainotti, and Antonio Pescapé. A tool for the generation of realistic network workload for emerging networking scenarios. *Comput. Netw.*, 56(15) :3531–3547, October 2012.
- [BKSP12] H. Bang, J. Kim, Y. Shin, and C. S. Park. Analysis of ont buffer and power management performances for xg-pon cyclic sleep mode. In *Global Communications Conference (GLOBECOM), 2012 IEEE*, pages 3116–3121, Dec 2012.
- [BRI] BRITE Integration - NS3.
- [CAP11] F. Cuomo, A. Abbagnale, and S. Papagna. Esol : Energy saving in the internet based on occurrence of links in routing paths. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium on*, pages 1–6, June 2011.

- [CCF⁺05] Matthew Caesar, Donald Caldwell, Nick Feamster, Jennifer Rexford, Aman Shaikh, and Jacobus van der Merwe. Design and implementation of a routing control platform. In *Proceedings of the 2Nd Conference on Symposium on Networked Systems Design & Implementation - Volume 2*, NSDI'05, pages 15–28, Berkeley, CA, USA, 2005. USENIX Association.
- [CCHW11] H. W. Chu, C. C. Cheung, K. H. Ho, and N. Wang. Green mpls traffic engineering. In *Australasian Telecommunication Networks and Applications Conference (ATNAC), 2011*, pages 1–4, Nov 2011.
- [CCN11] Zheng Cai, Alan L. Cox, and T. S. Eugene Ng. Maestro : A System for Scalable OpenFlow Control. Technical report, Rice University, 2011.
- [CCPM12] Francesca Cuomo, Antonio Cianfrani, Marco Polverini, and Daniele Mangione. Network pruning for energy saving in the internet. *Comput. Netw.*, 56(10) :2355–2367, July 2012.
- [CEL⁺10] A. Cianfrani, V. Eramo, M. Listanti, M. Marazza, and E. Vittorini. An energy saving routing algorithm for a green ospf protocol. In *INFOCOM IEEE Conference on Computer Communications Workshops , 2010*, pages 1–5, March 2010.
- [CELP11] A. Cianfrani, V. Eramo, M. Listanti, and M. Polverini. An ospf enhancement for energy saving in ip networks. In *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*, pages 325–330, April 2011.
- [CGL14] R. Carpa, O. Glămk, and L. Lefevre. Segment routing based traffic engineering for energy efficient backbone networks. In *2014 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, pages 1–6, Dec 2014.
- [CGM14] M. R. Celenlioglu, S. B. Goger, and H. A. Mantar. An sdn-based energy-aware routing model for intra-domain networks. In *Software, Telecommunications and*

- Computer Networks (SoftCOM), 2014 22nd International Conference on*, pages 61–66, Sept 2014.
- [CGNG04] Kenneth J. Christensen, Chamara Gunaratne, Bruce Nordman, and Alan D. George. The next frontier for communications networks : Power management. *Comput. Commun.*, 27(18) :1758–1770, December 2004.
- [Che00] T. M. Chen. Evolution to the programmable internet. *IEEE Communications Magazine*, 38(3) :124–128, Mar 2000.
- [cis] Cisco SG300-10 10-Port Gigabit Managed Switch.
- [Cis07] Cisco Data Center Infrastructure 2.5 Design Guide. http://www.cisco.com/application/pdf/en/us/guest/netsol/ns107/c649/ccmigration_09186a008073377d.pdf, 2007. [Online ; accessed 22-March-2016].
- [cis15] Cisco Visual Networking Index : Forecast and Methodology, 2014â “2019. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html, 2015. [Online ; accessed 22-March-2016].
- [CKY11] A. R. Curtis, W. Kim, and P. Yalagandula. Mahout : Low-overhead datacenter traffic management using end-host-based elephant detection. In *INFOCOM, 2011 Proceedings IEEE*, pages 1629–1637, April 2011.
- [CMN12] L. Chiaraviglio, M. Mellia, and F. Neri. Minimizing isp network energy cost : Formulation and solutions. *IEEE/ACM Transactions on Networking*, 20(2) :463–476, April 2012.
- [CRN⁺10] K. Christensen, P. Reviriego, B. Nordman, M. Bennett, M. Mostowfi, and J. A. Maestro. Ieee 802.3az : the road to energy efficient ethernet. *IEEE Communications Magazine*, 48(11) :50–56, November 2010.

- [CSB92] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen. Low-power cmos digital design. *IEEE Journal of Solid-State Circuits*, 27(4) :473–484, Apr 1992.
- [Dli] DGS-1100 Series, Gigabit Smart Managed Switches.
- [eee10] IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 3 : CSMA/CD Access Method and Physical Layer Specifications Amendment 5 : Media Access Control Parameters, Physical Layers, and Management Parameters for Energy-Efficient Ethernet. *IEEE Std 802.3az-2010 (Amendment to IEEE Std 802.3-2008)*, pages 1–302, Oct 2010.
- [ELCM13] V. Eramo, M. Listanti, A. Cianfrani, and E. Miucci. Evaluation of power saving in an mpls/ip network hosting a virtual router layer of a single service provider. In *ICT Convergence (ICTC), 2013 International Conference on*, pages 12–17, Oct 2013.
- [EPV04] D. Emma, A. Pescapé, and G. Ventre. Analysis and experimentation of an open distributed platform for synthetic traffic generation. In *Distributed Computing Systems, 2004. FTDCS 2004. Proceedings. 10th IEEE International Workshop on Future Trends of*, pages 277–283, May 2004.
- [FM15] Clarence Filsfils and Kris Michielsen. Segment Routing Introduction, Nov 2015.
- [FPD⁺16] C. Filsfils, S. Previdi, B. Decraene, S. Litkowski, and R. Shakir. Segment Routing Architecture draft-ietf-spring-segment-routing-08, May 2016.
- [FWM⁺14] F. Francois, N. Wang, K. Moessner, S. Georgoulas, and R. de Oliveira Schmidt. Leveraging mpls backup paths for distributed energy-aware traffic engineering. *IEEE Transactions on Network and Service Management*, 11(2) :235–249, June 2014.
- [GCN05] Chamara Gunaratne, Ken Christensen, and Bruce Nordman. Managing energy

- consumption costs in desktop pcs and lan switches with proxying, split tcp connections, and scaling of link speed. *Int. J. Netw. Manag.*, 15(5) :297–310, September 2005.
- [GCNS08] C. Gunaratne, K. Christensen, B. Nordman, and S. Suen. Reducing the energy consumption of ethernet with adaptive link rate (alr). *IEEE Transactions on Computers*, 57(4) :448–461, April 2008.
- [GCS06] C. Gunaratne, K. Christensen, and S. W. Suen. Ngl02-2 : Ethernet adaptive link rate (alr) : Analysis of a buffer threshold policy. In *Global Telecommunications Conference, 2006. GLOBECOM 06. IEEE*, pages 1–6, Nov 2006.
- [GGS04] M. Gupta, S. Grover, and S. Singh. A feasibility study for power management in lan switches. In *Network Protocols, 2004. ICNP 2004. Proceedings of the 12th IEEE International Conference on*, pages 361–371, Oct 2004.
- [GLL⁺09] Chuanxiong Guo, Guohan Lu, Dan Li, Haitao Wu, Xuan Zhang, Yunfeng Shi, Chen Tian, Yongguang Zhang, and Songwu Lu. Bcube : A high performance, server-centric network architecture for modular data centers. In *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication*, SIGCOMM '09, pages 63–74, New York, NY, USA, 2009. ACM.
- [GMP14] F. Giroire, J. Moulierac, and T. K. Phan. Optimizing rule placement in software-defined networks for energy-aware routing. In *Global Communications Conference (GLOBECOM), 2014 IEEE*, pages 2523–2529, Dec 2014.
- [Gre15] GreenTouch Final Results from Green Meter Research Study : Reducing the Net Energy Consumption in Communications Networks by up to 90% by 2020. http://www.ourenergypolicy.org/wp-content/uploads/2015/06/GreenTouch_Green_Meter_Final_Results_18_June_2015.pdf, 2015. [Online; accessed 22-March-2016].

- [GS03] Maruti Gupta and Suresh Singh. Greening of the internet. In *Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, SIGCOMM '03, pages 19–26, New York, NY, USA, 2003. ACM.
- [GS07a] M. Gupta and S. Singh. Dynamic ethernet link shutdown for energy conservation on ethernet links. In *Communications, 2007. ICC 07. IEEE International Conference on*, pages 6156–6161, June 2007.
- [GS07b] M. Gupta and S. Singh. Using low-power modes for energy conservation in ethernet lans. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 2451–2455, May 2007.
- [GS09] Erol Gelenbe and Simone Silvestri. *Quality of Service in Heterogeneous Networks : 6th International ICST Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness, QShine 2009 and 3rd International Workshop on Advanced Architectures and Algorithms for Internet Delivery and Applications, AAA-IDEA 2009, Las Palmas, Gran Canaria, November 23-25, 2009 Proceedings*, chapter Optimisation of Power Consumption in Wired Packet Networks, pages 717–729. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [HARPFVLG11] Sergio Herrería-Alonso, Miguel Rodríguez-Pérez, Manuel Fernández-Veiga, and Cándido López-García. Opportunistic power saving algorithms for ethernet devices. *Comput. Netw.*, 55(9) :2051–2064, June 2011.
- [HJP15] M. Hussein, G. Jakllari, and B. Paillassa. Frugal topologies for saving energy in ip networks. In *Local Computer Networks (LCN), 2015 IEEE 40th Conference on*, pages 303–311, Oct 2015.
- [HM8] HM8115-2 Power Meter. https://www.rohde-schwarz.com/product/hm8115-productstartpage_63493-44099.html. [Online; accessed 22-March-2016].

- [HMA⁺99] M. Hicks, J. T. Moore, D. S. Alexander, C. A. Gunter, and S. M. Nettles. Planet : an active internetwork. In *INFOCOM 99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 3, pages 1124–1133 vol.3, Mar 1999.
- [HSM⁺10] Brandon Heller, Srini Seetharaman, Priya Mahadevan, Yiannis Yiakoumis, Puneet Sharma, Sujata Banerjee, and Nick McKeown. Elastictree : Saving energy in data center networks. In *Proceedings of the 7th USENIX Conference on Networked Systems Design and Implementation*, NSDI’10, pages 17–17, Berkeley, CA, USA, 2010. USENIX Association.
- [IEE12] Ieee standard for ethernet. *IEEE Std 802.3-2012 (Revision to IEEE Std 802.3-2008)*, pages 1–3747, Dec 2012.
- [JCN08] M. Jimeno, K. Christensen, and B. Nordman. A network connection proxy to enable hosts to sleep and save energy. In *Performance, Computing and Communications Conference, 2008. IPCCC 2008. IEEE International*, pages 101–110, Dec 2008.
- [KCG⁺10] Teemu Koponen, Martin Casado, Natasha Gude, Jeremy Stribling, Leon Poutievski, Min Zhu, Rajiv Ramanathan, Yuichiro Iwata, Hiroaki Inoue, Takayuki Hama, and Scott Shenker. Onix : A distributed control platform for large-scale production networks. In *Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation*, OSDI’10, pages 351–364, Berkeley, CA, USA, 2010. USENIX Association.
- [KZY⁺14] D. A. Khotimsky, D. Zhang, L. Yuan, R. O. C. Hirafuji, and D. R. Campelo. Unifying sleep and doze modes for energy-efficient pon systems. *IEEE Communications Letters*, 18(4) :688–691, April 2014.
- [LGYN13] R. Liu, H. Gu, X. Yu, and X. Nian. Distributed flow scheduling in energy-aware data center networks. *IEEE Communications Letters*, 17(4) :801–804, April 2013.

- [LKB⁺14] C. Lange, D. Kosiankowski, A. Betker, H. Simon, N. Bayer, D. von Hugo, H. Lehmann, and A. Gladisch. Energy efficiency of load-adaptively operated telecommunication networks. *Journal of Lightwave Technology*, 32(4) :571–590, Feb 2014.
- [LLW⁺12] D. Li, Y. Li, J. Wu, S. Su, and J. Yu. Esm : Efficient and scalable data center multicast routing. *IEEE/ACM Transactions on Networking*, 20(3) :944–955, June 2012.
- [LSC15] Gongqi Lin, Sieteng Soh, and Kwan-Wu Chin. Energy-aware traffic engineering with reliability constraint. *Computer Communications*, 57 :115 – 128, 2015.
- [LSLC13] Gongqi Lin, S. Soh, M. Lazarescu, and Kwan-Wu Chin. Energy-aware two link-disjoint paths routing. In *High Performance Switching and Routing (HPSR), 2013 IEEE 14th International Conference on*, pages 103–108, July 2013.
- [M⁺13] A. Manzalini et al. Software-Defined Networks for Future Network and Services. <http://sites.ieee.org/sdn4fns/whitepaper/>, Nov 2013, Trento, Italy. [Online ; accessed 4-November-2016].
- [MC12] M. Mostowfi and K. Christensen. An energy-delay model for a packet coalescer. In *Southeastcon, 2012 Proceedings of IEEE*, pages 1–6, March 2012.
- [MLMB01] Alberto Medina, Anukool Lakhina, Ibrahim Matta, and John Byers. Brite : An approach to universal topology generation. In *Proceedings of the Ninth International Symposium in Modeling, Analysis and Simulation of Computer and Telecommunication Systems, MASCOTS '01*, pages 346–, Washington, DC, USA, 2001. IEEE Computer Society.
- [NPI⁺08] Sergiu Nedevschi, Lucian Popa, Gianluca Iannaccone, Sylvia Ratnasamy, and David Wetherall. Reducing network energy consumption via sleeping and rate-adaptation. In *Proceedings of the 5th USENIX Symposium on Networked Systems Design and*

- Implementation*, NSDI'08, pages 323–336, Berkeley, CA, USA, 2008. USENIX Association.
- [NS3] The Network Simulator - NS3.
- [ONF] Software-Defined Networking (SDN) Definition. <https://www.opennetworking.org/sdn-resources/sdn-definition>. [Online ; accessed 4-November-2016].
- [PCB01] B. Paillassa, X. Cregut, and A. Benarbia. Asa : the active service architecture for new internet. In *Communications, Computers and signal Processing, 2001. PACRIM. 2001 IEEE Pacific Rim Conference on*, volume 2, pages 489–492 vol.2, 2001.
- [PDJC13] B. Paillassa, R. Dhaou, G. Jakllari, and T. Chakadkit. Performance evaluation of energy efficient policies for ethernet switches. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2013 9th International*, pages 797–802, July 2013.
- [PN14] K. Pandiaraj and L. Nisha. A low power parity cam for asynchronous transfer mode(atm) switch application. In *Computer Communication and Informatics (IC-CCI), 2014 International Conference on*, pages 1–4, Jan 2014.
- [rfc98] rfc2328, ospf version 2, 1998.
- [RRCRS14] A. Ruiz-Rivera, K. W. Chin, R. Raad, and S. Soh. Hotpluz : A bgp-aware green traffic engineering approach. In *Communications (ICC), 2014 IEEE International Conference on*, pages 3721–3726, June 2014.
- [RRCS15] A. Ruiz-Rivera, K. W. Chin, and S. Soh. Greco : An energy aware controller association algorithm for software defined networks. *IEEE Communications Letters*, 19(4) :541–544, April 2015.

- [RRCsr15] Alejandro Ruiz-Rivera, Kwan-Wu Chin, Sieteng Soh, and Raad Raad. On the performance of online and offline green path establishment techniques. *EURASIP Journal on Wireless Communications and Networking*, 2015(1) :1–17, 2015.
- [SCvV⁺12] D. Suvakovic, H. Chow, D. van Veen, J. Galaro, B. Farah, N. P. Anthapadmanabhan, P. Vetter, A. Dupas, and R. Boislaigue. Low energy bit-interleaving downstream protocol for passive optical networks. In *Online Conference on Green Communications (GreenCom), 2012 IEEE*, pages 26–31, Sept 2012.
- [SGRO⁺08] K. Sabhanatarajan, A. Gordon-Ross, M. Oden, M. Navada, and A. George. Smart-nics : Power proxying for reduced power consumption in network edge devices. In *Symposium on VLSI, 2008. ISVLSI '08. IEEE Computer Society Annual*, pages 75–80, April 2008.
- [SLX10] Yunfei Shang, Dan Li, and Mingwei Xu. Energy-aware routing in data center network. In *Proceedings of the First ACM SIGCOMM Workshop on Green Networking, Green Networking '10*, pages 1–8, New York, NY, USA, 2010. ACM.
- [SLX⁺12] Meng Shen, Hongying Liu, Ke Xu, Ning Wang, and Yifeng Zhong. *NETWORKING 2012 : 11th International IFIP TC 6 Networking Conference, Prague, Czech Republic, May 21-25, 2012, Proceedings, Part I*, chapter Routing On Demand : Toward the Energy-Aware Traffic Engineering with OSPF, pages 232–246. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [SS13] Stefan Schmid and Jukka Suomela. Exploiting locality in distributed sdn control. In *Proceedings of the Second ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking, HotSDN '13*, pages 121–126, New York, NY, USA, 2013. ACM.
- [TE02] Rfc3272, overview and principles of internet te, 2002.
- [Tro08] powersaving modes for GPONand VDSL2. http://www.ieee802.org/3/10GEPON_study/email/pdfV3kikU0bA1.pdf, 2008. [Online ; accessed 12-May-2016].

- [TSGR08] R. Teixeira, A. Shaikh, T. G. Griffin, and J. Rexford. Impact of hot-potato routing changes in ip networks. *IEEE/ACM Transactions on Networking*, 16(6) :1295–1307, Dec 2008.
- [VHIV⁺12] Ward Van Heddeghem, Filip Idzikowski, Willem Vereecken, Didier Colle, Mario Pickavet, and Piet Demeester. Power consumption modeling in optical multilayer networks. *Photonic Netw. Commun.*, 24(2) :86–102, October 2012.
- [VHLL⁺14] Ward Van Heddeghem, Sofie Lambert, Bart Lannoo, Didier Colle, Mario Pickavet, and Piet Demeester. Trends in worldwide ict electricity consumption from 2007 to 2012. *Comput. Commun.*, 50 :64–76, September 2014.
- [WHPH08] N. Wang, K. H. Ho, G. Pavlou, and M. Howarth. An overview of routing optimization for internet traffic engineering. *IEEE Communications Surveys Tutorials*, 10(1) :36–56, First 2008.
- [WJG⁺14] R. Wang, Z. Jiang, S. Gao, W. Yang, Y. Xia, and M. Zhu. Energy-aware routing algorithms in software-defined networks. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2014 IEEE 15th International Symposium on a*, pages 1–6, June 2014.
- [WJZW12] J. Wu, Q. Jia, J. Zhao, and X. Wang. Minimizing the energy cost of servers in data center networks. In *Cloud and Green Computing (CGC), 2012 Second International Conference on*, pages 311–318, Nov 2012.
- [WMG10] D. Wang, J. McNair, and A. George. A smart-nic-based power-proxy solution for reduced power consumption during instant messaging. In *Green Technologies Conference, 2010 IEEE*, pages 1–10, April 2010.
- [XG-14] 10-Gigabit-capable passive optical networks (XG-PON) : Transmission convergence (TC) layer Specification. *ITU-T G.987.3*, Jan 2014.
- [Yin] Yin Zhangâs Abilene TM.

- [ZYLZ10] M. Zhang, C. Yi, B. Liu, and B. Zhang. Greente : Power-aware traffic engineering. In *Network Protocols (ICNP), 2010 18th IEEE International Conference on*, pages 21–30, Oct 2010.